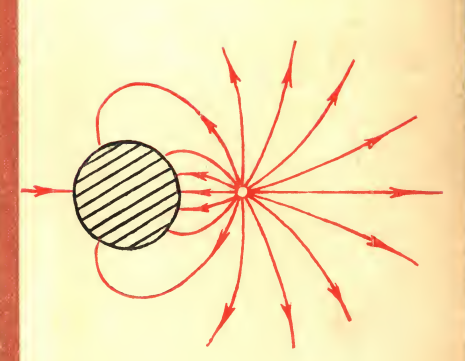
Е. И. БУТИКОВ, А. А. БЫКОВ, А. С. КОНДРАТЬЕВ

ФИЗИКА

ДЛЯ ПОСТУПАЮЩИХ В ВУЗЫ



Е. И. БУТИКОВ, А. А. БЫКОВ, А. С. КОНДРАТЬЕВ

ФИЗИКА

ДЛЯ ПОСТУПАЮЩИХ В ВУЗЫ

ИЗДАНИЕ ВТОРОЕ, ИСПРАВЛЕННОЕ

Допущено Министерством высшего и среднего специального образования СССР в качестве учебного пособия для слушателей подготовительных отделений высших учебных заведений

МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
1982

22.3
Б93
УДК 530.1

Бутиков Е. И., Быков А. А., Кондратьев А. С. Физика для поступающих в вузы: Учебное пособие.— 2-е изд., испр.— М.: Наука. Главная редакция физико-математической литературы, 1982.— 608 с.

Задача книги — способствовать развитию более широкого кругозора, навыков физического мышления и глубокого понимания основных физических законов, а также стимулировать интерес к предмету. Большое внимание уделено разбору конкретных физических задач и примеров. Используемый математический аппарат полностью соответствует школьной программе. В новом издании исправлены опечатки и отдельные неточности предыдущего издания, выходившего в 1978 г.

Рис. 319.

Евгений Иванович Бутиков
Александр Александрович Быков
Александр Сергеевич Кондратьев

ФИЗИКА

для поступающих в вузы

Редактор Б. С. Беликов.
Техн. редактор С. Я. Шкляр. Корректор Т. С. Плетнева.

ИБ № 12135

Печать с матриц. Подписано к печати 17.12.81. Формат $84\times108^1/_{32}$. Бумага тип. № 2. Литературная гарнитура. Высокая печать. Условн. печ. л. 31,92. Уч.-изд. л. 31,7. Тираж 300 000 экз. (1-й завод 1—200 000 экз). Заказ № 3565. Цена 1 р. 20 к.

Издательство «Наука»
Главная редакция физико-математической литературы
117071, Москва, В-71, Ленинский проспект, 15

Ордена Октябрьской Революции и ордена Трудового Красного Знамени Первая Образцовая типография имени А. А. Жданова Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли. Москва, М-54, Валовая, 28

Б $\frac{1704010000-013}{053(02)-82}$ 83-82

ПРЕДИСЛОВИЕ

Книга представляет собой пособие по курсу физики средней школы. Особое внимание в ней уделяется тем вопросам, которые по тем или иным причинам не изложены в школьном учебнике или изложены там недостаточно глубоко и подробно. Целый ряд вопросов, изложенных в школьном учебнике физики, в этой книге освещается с несколько иной точки зрения. Подбор материала осуществлен таким образом, чтобы способствовать развитию у читателя широкого кругозора и глубокого понимания основных физических законов.

Для чтения данной книги необходимо знакомство с материалом школьного учебника, ибо авторы стремились избегать его дублирования. В книге излагается не только теоретический материал: разъяснение сути фундаментальных физических законов проводится путем разбора большого числа оригинальных примеров и задач.

Объем знаний по физике, сообщаемых учащимся в средней школе, достаточно велик и в принципе позволяет рассматривать весьма сложные вопросы при анализе физических явлений. Однако опыт проведения конкурсных экзаменов в вузах показывает, что у некоторой части абитуриентов имеется разрыв между уровнем знаний фактического материала и умением применять эти знания на практике.

Предлагаемая книга поможет учащимся ликвидировать этот разрыв. На многочисленных примерах показывается, насколько эффективными могут быть получаемые в средней школе знания по физике при анализе даже очень сложных вопросов и задач при условии действительно глубокого понимания сути физических законов. Часто обсуждение этих вопросов не требует существенного расширения круга используемых в средней школе понятий.

Степень глубины физического понимания характеризуется умением применять для анализа различных явлений наиболее общие, фундаментальные законы. При анализе конкретных примеров и задач в разных разделах книги показывается, как, например, применение закона сохранения энергии позволяет решить задачу проще, взглянуть на нее с более общих позиций и, что особенно важно, часто дает возможность найти ответ на некоторые вопросы, касающиеся тех явлений, для которых нам неизвестны описывающие их конкретные законы. Для глубокого понимания физики необходимо четкое осознание степени общности различных физических законов, границ их применимости, их места в общей физической картине мира.

Книга не перегружена формулами. Там, где это возможно и не идет в ущерб строгости изложения, авторы стремились максимально использовать качественные соображения. По мнению одного из величайших физиков Э. Ферми, «...физическая сущность действительно понимаемого вопроса может быть объяснена без помощи сложных формул» *). В умении объяснить сущность вопроса «на пальцах» и заключается истинное понимание уравнений, выражающих физические законы.

Используемый в книге математический аппарат соответствует современной школьной программе по математике, содержащей элементы анализа и дающей учащимся представление об идейной стороне дифференциального и интегрального исчисления. Однако вычислительные процедуры, требующие наличия определенных навыков дифференцирования и интегрирования, за редким исключением, не используются. Напротив, в ряде примеров показывается, как глубокое понимание физики изучаемых явлений позволяет взглянуть на них с несколько иной точки зрения и обойти при этом вычислительные математические трудности.

Книга предназначена в основном для чтения после изучения школьного курса физики и математики. Однако ее можно читать и параллельно с изучением соответствующего материала в школе. При этом следует учитывать, что, как правило, в каждом параграфе материал расположен по

---

*) *Б. М. Понтекорво*, Энрико Ферми, «Знание», 1971.

нарастающей трудности. Поэтому при первом чтении можно не стремиться «пробиться» до конца каждого параграфа, а остановиться на том месте, где читатель почувствует, что для понимания дальнейшего материала у него еще нет необходимых знаний.

Преподавание физики в средней школе по необходимости ведется таким образом, что материал излагается строго последовательно. При этом у учащихся нет представления ни о других, еще не изученных разделах физики, ни о связях изучаемого материала с этими разделами. Одной из характерных особенностей этой книги является то, что в ней физика рассматривается как единая наука, все разделы которой взаимно связаны. Это предполагает у читателя наличие определенных представлений о содержании всех разделов физики. Изложение построено таким образом, чтобы учащемуся в дальнейшем, при изучении физики в высшей школе, не пришлось переучиваться. Разумеется, при более глубоком изучении предмета взгляд на некоторые вещи может существенно измениться, однако ничто из того, что читатель узнает из этой книги, не будет при этом объявлено неверным.

Изложенный в книге материал использовался на занятиях в физико-математической школе-интернате при Ленинградском университете. Учащиеся этой школы в течение многих лет систематически добиваются успехов на всесоюзных и международных олимпиадах школьников по физике и успешно выдерживают вступительные экзамены в ведущие вузы страны.

Книга может принести пользу всем, кто интересуется физикой и имеет склонность к изучению точных наук. Ее можно использовать на занятиях в школьных физических кружках, на факультативных занятиях по физике. Она, несомненно, окажется интересной для учителей физики средних школ, профессионально-технических училищ и техникумов. Книга будет полезна при подготовке к конкурсным экзаменам в высшие учебные заведения. По целому ряду вопросов она может представить интерес и для студентов вузов.

Авторы надеются, что книга будет с интересом прочитана всеми, кто так или иначе связан или будет связан с физикой и техникой в своей профессиональной деятельности.

Книга содержит следующие разделы: 1. Механика. 2. Молекулярная физика и термодинамика. 3. Электричество и магнетизм. 4. Колебания и волны. 5. Оптика. Теория относительности. 6. Введение в квантовую физику. Такая последовательность изложения материала соответствует программе по физике средней школы и программе вступительных экзаменов в вуз. В данной книге отсутствует раздел «Физика атомного ядра и элементарных частиц». Это связано главным образом с тем, что, по мнению авторов, материал этого раздела с достаточной полнотой представлен в школьном учебнике. Некоторые из относящихся сюда вопросов затронуты в других разделах книги.

Авторы считают своим приятным долгом выразить благодарность Б. Б. Буховцеву и В. И. Николаеву, взявшим на себя труд по рецензированию рукописи, и редактору Б. С. Беликову. Их доброжелательные замечания во многом способствовали улучшению книги. Авторы признательны Ю. Л. Климонтовичу, чья инициатива сыграла немалую роль в появлении этой книги.

Е. И. Бутиков
А. А. Быков
А. С. Кондратьев

К ЧИТАТЕЛЮ

Это обращение адресовано тем читателям, которые решили готовиться к вступительному экзамену по физике в вуз с помощью этой книги.

Вступительный экзамен в вуз в наше время не сводится к формальному ответу на вопросы экзаменационного билета, а зачастую представляет собой беседу с абитуриентом, во время которой выясняются не только его фактические знания, а главным образом круг его интересов, степень его подготовленности для обучения в данном вузе. Поэтому подготовка к экзамену не должна сводиться только к повторению изученного в школе материала, а должна обязательно углублять полученные в школе знания и расширять кругозор. При этом не следует забывать, что подготовка к экзамену необходима не только для того, чтобы преодолеть этот барьер, но и для укрепления фундамента, на котором будет построено дальнейшее обучение в вузе.

Приступать к чтению этой книги следует, располагая достаточным запасом времени. Книга насыщена материалом, поэтому ее чтение потребует временами от читателя некоторых усилий. Во всяком случае читать ее следует с карандашом в руке, обязательно воспроизводя некоторые опущенные выкладки (как правило, в тексте приводятся необходимые указания, как делать эти выкладки). Не пренебрегайте этим советом, так как многие выкладки в книге не приведены потому, что именно самостоятельное их воспроизведение читателем способствует наиболее глубокому пониманию и эффективному усвоению материала.

Обычно в каждом параграфе материал расположен по нарастающей трудности. Поэтому при первом чтении книги можно не стремиться обязательно дочитать параграф до конца. Затруднения могут быть связаны с недостаточным знанием материала школьного учебника по этому вопросу. Тогда целесообразно открыть школьный учебник и освежить

в памяти соответствующий материал. Причиной затруднений может быть и недостаточная математическая подготовка. В этом случае важно не падать духом, а проявить определенную настойчивость,— в учении никогда не бывает легко!

Изучение этой книги должно помочь читателю составить представление о предмете в целом и об отдельных его разделах. Умение уверенно ориентироваться в обширном материале особенно важно потому, что абитуриенту известна только программа экзамена. Конкретные вопросы, входящие в экзаменационные билеты, могут быть сформулированы по-разному, и для успешного ответа абитуриенту прежде всего необходимо самому сделать правильный отбор нужного материала. Поэтому, встретив в этой книге не изучавшийся в школе материал, не торопитесь отбрасывать его в сторону. Разобравшись в нем, можно по-новому взглянуть на некоторые хорошо известные вещи и обнаружить взаимосвязь между явлениями, ранее казавшимися совершенно независимыми. Читая книгу, задумывайтесь над тем, на какие вопросы можно ответить, используя изучаемый материал.

В книге подробно рассмотрено большое число конкретных примеров, иллюстрирующих содержание фундаментальных физических законов и применение этих законов. Но авторы рекомендуют не ограничиваться разбором только приведенных примеров. Обязательно нужно решать задачи. Только самостоятельно решая задачи, можно достичь того высокого уровня понимания физических законов, при котором появится возможность отвечать на новые вопросы, относящиеся к изучаемой области.

1. МЕХАНИКА

КИНЕМАТИКА

§ 1. Пространство и время. Системы отсчета. Основные понятия кинематики материальной точки

Механическое движение — это перемещение изучаемого тела относительно другого тела, принимаемого за неподвижное. Одно и то же механическое движение может выглядеть совершенно по-разному в зависимости от того, какое тело при изучении этого движения считается неподвижным. Так, например, с точки зрения наблюдателя, находящегося в вагоне поезда, выскользнувший из рук предмет падает вертикально вниз, а для наблюдателя на земле это же движение

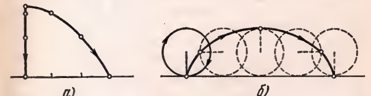


Рис. 1.1. Траектории свободно падающего тела (*а*) и точки на ободе колеса (*б*) в разных системах отсчета.

происходит по параболической траектории. Точка на ободе колеса вагона движется по окружности для первого из этих наблюдателей, но это же движение с точки зрения второго наблюдателя происходит по замысловатой кривой — циклоиде (рис. 1.1). В этом заключается простейшее проявление относительного характера механического движения.

Механическое движение происходит в пространстве и во времени. Понятия пространства и времени относятся к числу фундаментальных понятий, которые невозможно определить через какие-то более простые понятия. Но в физике

более важными являются не формальные определения, а конкретные свойства пространства и времени, проявляющиеся при протекании тех или иных физических процессов. И эти конкретные свойства познаются опытным путем.

Физика — точная наука, т. е. в ее основе лежит изучение не только качественных, но и количественных соотношений, поэтому любой физический опыт связан с измерениями. По словам Д. И. Менделеева, наука начинается с тех пор, как начинают измерять.

Для изучения механического движения, происходящего в пространстве и во времени, нужно прежде всего уметь измерять промежутки времени и расстояния.

В различных областях физики и техники используются разные способы измерения расстояний и промежутков времени. Любое измерение сводится к сравнению измеряемой величины с другой величиной, принятой за единичную, т. е. с эталоном. Выбор эталона и способа сравнения измеряемой величины с эталоном может быть сделан по-разному.

В настоящее время в качестве эталона длины принят метр — длина, равная 1 650 763,73 длин световых волн в вакууме определенной оранжевой линии в спектре излучения атома криптона-86.

В качестве эталона времени принята секунда — промежуток времени, равный 9 192 631 770 периодам электромагнитного излучения, соответствующего определенному переходу в атоме цезия-133 в отсутствие внешних полей.

Использовавшийся ранее эталон метра в виде стержня из сплава платины и иридия оказался недостаточно надежным, ибо, как и все твердые тела, подвержен внешним влияниям и не может быть воспроизведен в случае его утраты. Точно так же не удовлетворяет современным требованиям эталон времени, основанный на использовании «астрономических часов», т. е. суточного или орбитального вращения Земли. Не исключено, что технический прогресс приведет к такой точности измерений, что и существующие эталоны придется заменить на новые, более совершенные.

Существуют различные способы сравнения измеряемой величины с эталоном. Например, при измерении длины возможно непосредственное сравнение измеряемого отрезка с промежуточными эталонами — жесткими линейками. Такой способ основан на свойстве твердых тел сохранять до известной степени неизменными свою форму и размеры.

Другой возможный способ измерения длины — триангуляция (рис. 1.2). В этом способе непосредственно измеряют длину «базы» AB, на концах которой измеряют углы α и β в направлении на объект C. Затем искомое расстояние AC (или BC) рассчитывается по формулам тригонометрии. В основе этого способа лежит подтверждаемая опытом гипотеза о том, что световые лучи, приходящие от объекта C к точкам A и B, подчиняются тем же аксиомам геометрии Евклида, что и геометрические прямые линии.

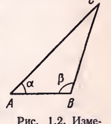


Рис. 1.2. Измерение расстояний методом триангуляции.

Еще один возможный способ измерения расстояний — радиолокация (или светолокация). При этом измеряется время прохождения электромагнитного сигнала до объекта и обратно и предполагается, что сигнал распространяется туда и обратно с одинаковой скоростью.

Вопрос о том, согласованы ли между собой разные способы измерения расстояний, т. е. дают ли они одно и то же или нет, решается только на опыте.

Измерение промежутков времени также возможно не только путем непосредственного сравнения с эталоном, но и путем использования различных косвенных методов, основанных на физических явлениях разной природы. Например, возможно использование периодических механических колебаний, как это делается в обычных маятниковых и кварцевых часах. Для измерения больших промежутков времени, например, в геологии и археологии, используется явление радиоактивного распада.

Для описания механического движения, как и любых других физических процессов, протекающих в пространстве и во времени, используются системы отсчета. Под системой отсчета понимают тело или систему тел, относительно которых рассматривается изучаемое движение, вместе с совокупностью связанных с этими телами приборов для измерения расстояний и промежутков времени.

Одно и то же физическое явление можно рассматривать в разных системах отсчета. Хотя изучаемое явление в разных системах отсчета может выглядеть по-разному, но длины тел и промежутки времени, как показывает опыт, при движении со скоростями, малыми по сравнению со

скоростью света, являются абсолютными, т. е. не зависят от того, в какой системе отсчета они измеряются.

Кинематика описывает механическое движение, отвлекаясь от физических причин, его вызывающих. Выбор системы отсчета в кинематике определяется исключительно соображениями удобства при математическом описании. Никаких принципиальных преимуществ у одной системы отсчета по сравнению с другой в кинематике нет. Преимущества определенного класса систем отсчета — инерциальных систем — выявляются только в динамике.

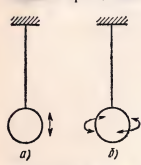


Рис. 1.3. Вертикальные (*а*) и крутильные (*б*) колебания тяжелого шара на упругой проволоке.

Из-за сложности физического мира, изучая реальное явление, мы всегда вынуждены упрощать его и вместо самого явления рассматривать некоторую идеализированную его модель, стремясь к тому, чтобы в выбранной модели сохранить самые характерные, наиболее важные черты явления. По образному выражению Я. И. Френкеля, физики фактически рассматривают не само явление, а карикатуру на него, и успех зависит от того, насколько удачна эта карикатура.

Простейшей моделью в кинематике является материальная точка. Материальной точкой считается любое тело, размеры которого в рассматриваемом явлении несущественны и их можно не принимать во внимание. Одно и то же тело в одних условиях можно считать материальной точкой, а в других — нельзя. Например, тяжелый шар, подвешенный на упругой проволоке, можно считать материальной точкой при изучении вертикальных колебаний (рис. 1.3, *а*) и нельзя — при изучении крутильных колебаний вокруг вертикальной оси (рис. 1.3, *б*). Таким образом, используя модель материальной точки, мы идеализируем не столько свойства самого тела, сколько условия его движения.

Положение материальной точки в некоторой системе отсчета можно определить, задавая ее радиус-вектор $\boldsymbol{r}$. Если связать с системой отсчета координатные оси, например, декартовой прямоугольной системы координат, то

задаиие радиус-вектора $\boldsymbol{r}$ эквивалеитио задаиию трех координат x, y, z — проекций радиус-вектора иа выбранные осн. Двнжеиие материальной точки математически описаио полиостью, если известеи ее радиус-вектор как функция времени $\boldsymbol{r}(t)$, т. е. известны три скалярные функцин $x(t)$, $y(t)$ и $z(t)$.

Измеиеиие радиус-вектора за иекоторый промежуток времеии $\Delta t = t_2 - t_1$, равное $\boldsymbol{r}(t_2) - \boldsymbol{r}(t_1)$, называют перемещеинем $\Delta \boldsymbol{r}$ за время Δt. Линия, которую описывает при этом конец раднус-вектора $\boldsymbol{r}(t)$, называется траекторней материальной точкн, а длнна этой ліннин — пройдеиным за время Δt путем ΔS (рнс. 1.4).

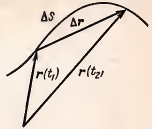


Рис. 1.4. Траектория, пройдеиный путь ΔS и перемещеиие $\Delta \boldsymbol{r}$ материальной точки.

Отиошенне вектора перемещення $\Delta \boldsymbol{r}$ к промежутку временн Δt иазывается средней по времени скоростью движення: $\boldsymbol{v}_{ср} = \Delta \boldsymbol{r}/\Delta t$. Если уменьшать зиачение промежутка времеин Δt, то отношеиие $\Delta \boldsymbol{r}/\Delta t$ будет стремиться к пределу, который называется мгиовенной скоростью (илн просто скоростью):

$$\boldsymbol{v} = \lim_{\Delta t \to 0} \frac{\Delta \boldsymbol{r}}{\Delta t} = \frac{d\boldsymbol{r}}{dt}. \qquad (1.1)$$

Таким образом, мгиовениая скорость представляет собой прonзводную раднус-вектора $\boldsymbol{r}$ по времеии. Наряду с прнведенным обозначеинем $d\boldsymbol{r}/dt$ для производной часто нспользуется другое обозначенне: $\boldsymbol{r}'$. Из самого определеиня скоростн следует, что в каждой точке вектор скоростн направлен по касательной к траекторни.

Часто нспользуется еще одна, ио уже скаляриая велнчнна, также называемая скоростью,— средний модуль скорости $\bar{v}$, определяемый как отношенне путн ΔS к промежутку времеии Δt, за который этот путь пройден: $\bar{v} = \Delta S/\Delta t$. Средняя скорость $\boldsymbol{v}_{ср} = \Delta \boldsymbol{r}/\Delta t$ н среднний модуль скорости $\bar{v}$, вообще говоря, характеризуют двнженне с разных сторои. Напрнмер, прн двнжеиии автомобиля по замкнутому пути вектор средней скорости $\boldsymbol{v}_{ср}$, вычнсляемый за полиое время двнжения, равеи иулю, иесмотря иа то, что

автомобиль прошел немалый путь, а средний модуль скорости $\bar{v}=\Delta S/\Delta t$ отличен от нуля. Однако при стремлении промежутка времени Δt к нулю величина вектора средней скорости $\boldsymbol{v}_{\text{ср}}$ приближается к среднему модулю скорости $\bar{v}$ и в пределе $\Delta t \to 0$ эти величины совпадают. Это как раз та

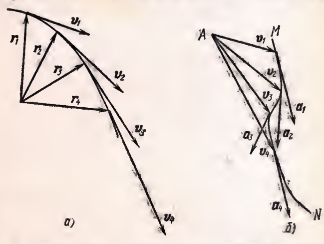


Рис. 1.5. Траектория (*а*) и годограф вектора скорости (*б*) материальной точки.

величина, которую показывает стрелка на спидометре автомобиля.

Вектор скорости характеризует быстроту изменения вектора перемещения материальной точки. Для характеристики быстроты изменения вектора скорости вводят ускорение. Средним за время Δt ускорением $\boldsymbol{a}_{\text{ср}}$ называется отношение приращения скорости $\Delta\boldsymbol{v}=\boldsymbol{v}(t_2)-\boldsymbol{v}(t_1)$ к промежутку времени Δt: $\boldsymbol{a}_{\text{ср}}=\Delta\boldsymbol{v}/\Delta t$. Предел этого отношения при $\Delta t \to 0$ называется мгновенным ускорением (или просто ускорением):

$$\boldsymbol{a}=\lim_{\Delta t\to 0}\frac{\Delta\boldsymbol{v}}{\Delta t}=\frac{d\boldsymbol{v}}{dt}. \quad (1.2)$$

Сравнивая формулы (1.1) и (1.2), можно отметить следующую формальную аналогию между скоростью и ускорением. Если скорость характеризует быстроту изменения радиус-вектора, то ускорение характеризует быстроту из-

менения вектора скорости. Рассмотрим эту аналогию подробнее. Пусть конец радиус-вектора описывает некоторую траекторию, показанную на рис. 1.5, *а*. В каждый момент времени вектор скорости направлен по касательной к траектории. Изобразим все векторы скорости $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ и т. д. так, чтобы они начинались в одной произвольной точке *A*

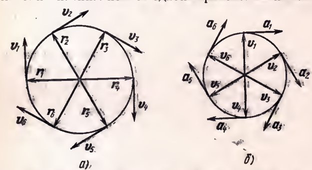


Рис. 1.6. Траектория (*а*) и годограф вектора скорости (*б*) при равномерном движении по окружности.

(рис. 1.5, *б*). При движении материальной точки по траектории конец вектора скорости на таком чертеже будет описывать кривую *MN*, называемую годографом вектора скорости. Используя такое определение, можно сказать, что сама траектория материальной точки является годографом ее радиус-вектора. Теперь легко сообразить, что вектор ускорения на рис. 1.5, *б* будет в каждой точке направлен по касательной к годографу вектора скорости *MN* подобно тому, как вектор скорости направлен по касательной к траектории на рис. 1.5, *а*.

Описанная аналогия может быть использована, например, для нахождения ускорения точки, равномерно движущейся по окружности (рис. 1.6, *а*). Годограф вектора скорости для такого движения показан на рис. 1.6, *б*. Пока материальная точка совершает один оборот по траектории, конец вектора скорости совершает один оборот по годографу. Величина скорости материальной точки связана с радиусом окружности *R* и периодом обращения *T* соотношением

$$v = \frac{2\pi R}{T}.$$

Аналогичное соотношение связывает величину ускорения a с радиусом годографа скорости v:

$$a = \frac{2\pi v}{T}.$$

Сравнивая эти формулы, получаем

$$a = \frac{v^2}{R}. \qquad (1.3)$$

Сравнивая рис. 1.6, *а* и *б*, убеждаемся, что вектор ускорения $\boldsymbol{a}$ в каждый момент времени направлен противоположно радиус-вектору материальной точки для этого же момента времени, т. е. ускорение $\boldsymbol{a}$ направлено к центру окружности, являющейся траекторией движения.

В рассматриваемом примере равномерного движения точки по окружности вектор скорости изменяется только по направлению, оставаясь неизменным по абсолютной величине, т. е. по модулю. Вектор ускорения при этом направлен перпендикулярно вектору скорости, т. е. по нормали к траектории. Так будет при движении с постоянной по величине скоростью по любой траектории. Если же скорость точки меняется и по величине, то у вектора ускорения кроме нормальной составляющей, направленной перпендикулярно скорости, будет еще составляющая, направленная по или против вектора скорости в зависимости от того, увеличивается или уменьшается скорость по величине. Величина нормальной составляющей ускорения определяется по-прежнему формулой (1.3), где под R понимается радиус кривизны траектории в данной точке, т. е. радиус окружности, дуга которой приблизительно совпадает с участком траектории вблизи рассматриваемой точки. Величина составляющей ускорения, параллельной скорости и характеризующей изменение величины скорости, равна производной от величины скорости по времени dv/dt.

§ 2. Кинематика движения в однородном поле

Простейший случай неравномерного движения — это движение с постоянным ускорением. Такое движение происходит в постоянном во времени однородном силовом поле. Примером такого поля может служить поле тяготения вблизи поверхности Земли при условии, что движение тела

происходит в области, линейные размеры которой малы по сравнению с радиусом Земли. Другой пример — действующее на заряженную частицу электрическое поле в пространстве между пластинами плоского конденсатора. Разумеется, движение тел в таких полях происходит с постоянным ускорением лишь при условии, что никакие другие силы, как, например, сопротивление воздуха, не действуют.

Уравнение движения тела, движущегося с постоянным ускорением $\boldsymbol{a}$, имеет вид

$$\boldsymbol{r} = \boldsymbol{r}_0 + \boldsymbol{v}_0 t + \frac{\boldsymbol{a}t^2}{2}. \qquad (2.1)$$

При одном и том же ускорении $\boldsymbol{a}$ движение может выглядеть совершенно по-разному в зависимости от того, каковы начальные условия — положение тела $\boldsymbol{r}_0$ в начальный момент времени и начальная скорость $\boldsymbol{v}_0$. Но во всех случаях это движение описывается одним и тем же уравнением (2.1). Такое движение происходит в одной плоскости, проходящей через векторы ускорения $\boldsymbol{a}$ и начальной скорости $\boldsymbol{v}_0$. Легко получить уравнение траектории движения.

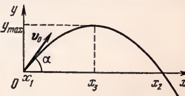


Рис. 2.1. Траектория движения материальной точки вблизи поверхности Земли.

Будем для определенности рассматривать движение вблизи поверхности Земли. Тогда ускорение $\boldsymbol{a}$ есть ускорение свободного падения $\boldsymbol{g}$ и, следовательно, траектория лежит в вертикальной плоскости. Введем в этой плоскости систему координат следующим образом: ось x направим горизонтально, а ось y — вертикально вверх. Спроектируем векторное уравнение (2.1) на оси x и y. Пусть α — угол между направлением начальной скорости и осью x. Тогда

$$x = x_0 + v_0 \cos\alpha \cdot t, \qquad y = y_0 + v_0 \sin\alpha \cdot t - \frac{gt^2}{2}. \qquad (2.2)$$

Обычно бывает удобно выбрать систему координат так, чтобы ее начало совпадало с исходной точкой траектории (рис. 2.1). Тогда $x_0 = y_0 = 0$ и уравнения (2.2) принимают вид

$$x = v_0 \cos\alpha \cdot t, \qquad y = v_0 \sin\alpha \cdot t - \frac{gt^2}{2}. \qquad (2.3)$$

Чтобы получить уравнение траектории $y=y(x)$, нужно исключить время из этих уравнений. Выражая t из первого уравнения и подставляя во второе, получим

$$y=x\operatorname{tg}\alpha-\frac{gx^2}{2v_0^2\cos^2\alpha}. \qquad (2.4)$$

Это уравнение параболы, проходящей через начало координат. Ее ветви направлены вниз, так как коэффициент при x^2 отрицателен. Используя хорошо известные свойства квадратного трехчлена, можно найти корни правой части выражения (2.4), ее наибольшее значение y_{max} и значение x, при котором достигается значение y_{max}. Корни трехчлена суть точки, в которых парабола (2.4) пересекает ось x. Один корень $x_1=0$ соответствует начальной точке траектории, второй $x_2=2(v_0^2/g)\sin\alpha\cdot\cos\alpha=(v_0^2/g)\sin 2\alpha$ дает дальность полета тела по горизонтали. Вершина параболы лежит на ее оси симметрии, т. е. посередине между корнями при $x=x_3=(v_0^2/2g)\sin 2\alpha$. Подставляя это значение x в (2.4), находим максимальную высоту подъема $y_{max}=(v_0^2/2g)\sin^2\alpha$. Если интересоваться тем, как будет меняться траектория при изменении направления начальной скорости, т. е. угла α, то удобнее преобразовать уравнение траектории (2.4) таким образом, чтобы оно содержало только какую-нибудь одну тригонометрическую функцию угла α. Для этого воспользуемся соотношением $1/\cos^2\alpha=1+\operatorname{tg}^2\alpha$. Подставляя его в (2.4), находим

$$y=x\operatorname{tg}\alpha-(1+\operatorname{tg}^2\alpha)\frac{gx^2}{2v_0^2}. \qquad (2.5)$$

Уравнение (2.5) описывает семейство параболических траекторий, зависящее от двух параметров: величины начальной скорости v_0 и угла α. Решение кинематических задач о свободном падении в однородном поле тяжести фактически сводится к исследованию этого семейства.

В качестве примера рассмотрим баллистическую задачу о стрельбе из ружья, пренебрегая сопротивлением воздуха. Прежде всего зададимся вопросом, как следует стрелять, чтобы попасть в цель, находящуюся на расстоянии S по горизонтали и на высоте h над горизонтальной плоскостью, проходящей через ружье (рис. 2.2). Стреляя в цель, мы можем менять наклон ствола ружья α, но, разумеется, мы не в силах менять величину начальной скорости v_0, так как она

зависит от варяда патронов и устройства ружья. Поэтому будем считать v_0 известной заданной величиной. Под каким же углом к горизонту следует направить ствол ружья?

Чтобы ответить на этот вопрос, потребуем, чтобы траектория, описываемая уравнением (2.5), проходила через цель, т. е. точку с координатами $x=S$, $y=h$:

$$h = S\,\mathrm{tg}\,\alpha - (1+\mathrm{tg}^2\,\alpha)\frac{gS^2}{2v_0^2}. \tag{2.6}$$

Это квадратное уравнение относительно $\mathrm{tg}\,\alpha$. Решая его, получаем для корней следующее выражение:

$$\mathrm{tg}\,\alpha = \frac{1}{gS}\left[v_0^2 \pm \sqrt{v_0^4 - g\,(gS^2 + 2v_0^2 h)}\right]. \tag{2.7}$$

Если дискриминант не отрицателен, т. е.

$$v_0^4 - g\,(gS^2 + 2v_0^2 h) \geqslant 0, \tag{2.8}$$

то уравнение имеет вещественные корни и, следовательно, при данной начальной скорости пули в цель попасть можно. Если при этом дискриминант положителен, т. е. уравнение (2.6) имеет два различных вещественных корня, то в цель можно попасть по двум различным траекториям. Траектория с меньшим значением угла α называется настильной, с бо́льшим — навесной. При равном нулю дискриминанте, когда корни (2.7) совпадают, в цель при данном значении начальной скорости можно попасть единственным образом. Если же дискриминант отрицателен, то уравнение (2.6) не имеет вещественных корней и в цель при данном значении v_0 попасть нельзя ни при каком значении угла α: ни одна из траекторий семейства (2.5) не «дотягивает» до этой цели. Отсюда ясно, что равенство нулю дискриминанта определяет ту минимальную начальную скорость $v_{0\,\min}$, при которой еще можно попасть в данную цель:

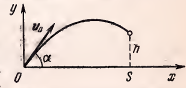


Рис. 2.2. Траектория пули при стрельбе в цель.

$$v_{0\,\min}^2 = g\left(h + \sqrt{h^2 + S^2}\right).$$

С другой стороны, при заданном значении v_0 равенство нулю дискриминанта определяет координаты наиболее удаленных

целей, в которые еще можно попасть, т. е. границу области, прострелнваемой из даииого ружья. Выражая из (2.8) h (в случае равеиства), находим

$$h=\frac{v_0^2}{2g}-\frac{gS^2}{2v_0^2}. \tag{2.9}$$

Эта формула определяет иаибольшую высоту цели, находящейся иа расстоянии S от ружья по горнзонталн, в которую еще можно попасть при даином v_0. С ее помощью легко

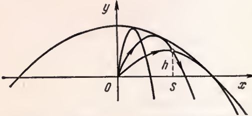


Рис. 2.3. Граннца простреливаемой области.

получить уравиеиие границы простреливаемой области, еслн заменить коорднаты определенной иаиболее удаленной цели S и h иа переменные велнчины x н y — коордииаты точек искомой границы:

$$y=\frac{v_0^2}{2g}-\frac{gx^2}{2v_0^2}. \tag{2.10}$$

Это уравиеиие параболы с вершиной при $x=0$, $y=v_0^2/2g$. Ее ветвн иаправлены вииз и пересекают горнзоитальную ось в точках $x=\pm v_0^2/g$ (рнс. 2.3). Все траекторнн с данным v_0 прн разиых значеннях α, т. е. семейство парабол (2.5), целнком лежат под этой граиицей, и каждая из траекторнй касается границы в одной точке. Другимн словами, граинца является огибающей для семейства таких траекторий. Через каждую цель, расположеиную инже границы, проходят две траекторнн, причем иавесная касается границы до попадаиня в цель.

Фактическн граница простреливаемой области представляет собой некоторую поверхность, а парабола (2.10) есть сечеиие этой поверхности вертнкальной плоскостью, проходящей через иачало коордииат. Вся поверхиость

может быть получена вращением параболы (2.10) вокруг оси y.

Если с самого начала интересоваться только границей простреливаемой области, то ее можно найти сразу с помощью уравнения траектории (2.5). Действительно, рассмотрим цели, находящиеся на одной вертикали, отстоящей от ружья на расстояние x, и найдем на этой вертикали самую высокую точку, в которую еще может попасть пуля. Эта точка, очевидно, принадлежит границе. Таким образом, задача сводится к нахождению максимума y при заданном x, т. е. максимума квадратного трехчлена (2.5) относительно $\operatorname{tg}\alpha$. Квадратный трехчлен имеет максимум при $\operatorname{tg}\alpha = v_0^2/gx$. Соответствующее максимальное значение получается подстановкой $\operatorname{tg}\alpha = v_0^2/gx$ в (2.5). Результат совпадает с формулой (2.10).

Полученные выше результаты, как нетрудно убедиться, содержат все хорошо известные частные случаи. Так, например, максимальная высота подъема $y_{\max} = v_0^2/2g$ получается из уравнения (2.10) при $x=0$, а наибольшая дальность полета пули по горизонтали при условии, что ружье и цель находятся на одной высоте, получается из (2.10) при $y=0$: $x_{\max} = v_0^2/g$.

Теперь несколько усложним задачу и вместо пули рассмотрим полет дробинок, имеющих одинаковые по величине, но несколько различающиеся по углу с горизонтом начальные скорости. Покажем, что при малом разбросе направлений начальных скоростей все дробинки в полете пройдут почти через одну и ту же точку. Пусть ствол ружья образует угол α с горизонтом (рис. 2.4). Будем характеризовать отклонение направления начальной скорости каждой дробинки от направления ствола углом δ, который будем считать малым: $\delta \ll \alpha$. Траектория дробинки, вылетевшей из ствола точно под углом α к горизонту (т. е. с $\delta=0$), определяется уравнением (2.5). Уравнение траектории дробинки, вылетевшей под углом $\alpha+\delta$, получается из (2.5) заменой α на $\alpha+\delta$:

$$y = x \operatorname{tg}(\alpha+\delta) - \frac{gx^2}{2v_0^2}\left[1 + \operatorname{tg}^2(\alpha+\delta)\right]. \qquad (2.11)$$

Наша задача — выяснить, пересекаются ли эти траектории. Если окажется, что пересекаются и при этом положение точки пересечения хотя бы приближенно не зависит от

величины δ, то это и будет означать, что траектории всех дробинок пройдут через эту точку. Такую точку можно назвать фокусом пучка частиц.

Итак, для нахождения положения фокуса нужно найти решение системы уравнений (2.5) и (2.11), кроме очевидного

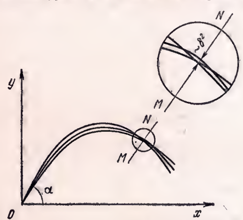


Рис. 2.4. Траектории дробинок при выстреле из ружья.

решения $x=y=0$, соответствующего точке вылета дробинок из ствола.

Преобразуем уравнение (2.11), воспользовавшись малостью угла δ. Для этого используем формулу для тангенса суммы двух углов и учтем, что для малых углов $\operatorname{tg}\delta\approx\delta$:

$$\operatorname{tg}(\alpha+\delta)\approx\frac{\operatorname{tg}\alpha+\delta}{1-\delta\operatorname{tg}\alpha}. \qquad (2.12)$$

Учитывая, что при $\gamma\ll 1$ справедливо соотношение $\frac{1}{1-\gamma}\approx 1+\gamma$ *), преобразуем (2.12), сохраняя только линейные по δ члены:

$$\operatorname{tg}(\alpha+\delta)\approx(\operatorname{tg}\alpha+\delta)(1+\delta\operatorname{tg}\alpha)\approx\operatorname{tg}\alpha+\delta(1+\operatorname{tg}^2\alpha). \qquad (2.13)$$

*) Действительно, умножая числитель и знаменатель дроби $\frac{1}{1-\gamma}$ на $1+\gamma$, получаем $\frac{1+\gamma}{1-\gamma^2}$. Пренебрегая γ^2 по сравнению с единицей, получаем приведенную формулу.

Подчеркнем, что в (2.13) отброшены члены порядка δ^2 и выше. Используя (2.13), получим, что в этом же приближении

$$1+\operatorname{tg}^2(\alpha+\delta)\approx 1+\operatorname{tg}^2\alpha+2\delta\operatorname{tg}\alpha\,(1+\operatorname{tg}^2\alpha). \quad (2.14)$$

Подставляя (2.13) и (2.14) в уравнение траектории (2.11), найдем

$$y=x\operatorname{tg}\alpha-\frac{gx^2}{2v_0^2}(1+\operatorname{tg}^2\alpha)+\\ +\delta\left[x(1+\operatorname{tg}^2\alpha)-\frac{gx^2}{v_0^2}\operatorname{tg}\alpha\,(1+\operatorname{tg}^2\alpha)\right]. \quad (2.15)$$

Рассматривая уравнения (2.5) и (2.15) как систему уравнений для нахождения координат фокуса x и y, видим, что положение фокуса, если он существует, не зависит от δ, ибо для его нахождения нужно просто приравнять нулю квадратную скобку в выражении (2.15):

$$x=\frac{v_0^2}{g\operatorname{tg}\alpha}. \quad (2.16)$$

Подставляя найденное значение x в уравнение (2.5), определяем y-координату фокуса:

$$y=\frac{v_0^2}{2g}(1-\operatorname{ctg}^2\alpha). \quad (2.17)$$

При нахождении координат фокуса мы пренебрегали в уравнении траектории всеми степенями δ выше первой, поэтому полученное свойство фокусировки (независимость координат точки пересечения разных траекторий от δ) является приближенным и справедливо с точностью до членов, линейных по δ. На самом деле пересечение различных траекторий происходит не в одной точке, а в некоторой малой области, линейные размеры которой в направлении, перпендикулярном к траектории, пропорциональны δ^2 (рис. 2.4).

Из формул (2.16) и (2.17) видно, что положение фокуса зависит от угла α, т. е. от наклона ствола ружья. Исключая из (2.16) и (2.17) угол α, получим геометрическое место фокусов при всевозможных значениях α:

$$y=\frac{v_0^2}{2g}-\frac{gx^2}{2v_0^2}.$$

Сравиивая это выражеиие с формулой (2.10), немедлеиио убеждаемся, что фокусировка пучка дробинок происходит как раз в точках граиицы достижимых при даииой иачальиой скорости целей. Это еще раз указывает иа приближеииый характер фокусировки, ибо выше было показано, что через каждую точку граиицы проходит только одиа траектория.

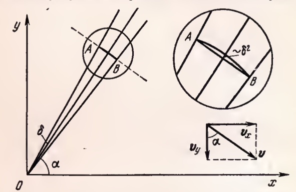


Рис. 2.5. Траектории дробинок в отсутствие тяготения.

На самом деле пересечеиие траекторий пучка частиц происходит ие в самой точке граиицы, а в иекоторой малой области вблизи этой точки, ио, коиечио, эта область целиком лежит ииже граиицы.

Свойство фокусировки пучка частиц можио устаиовить ие только с помощью уравиеиий траекторий, но и иепосредствеиио из уравиеиий движеиия.

Проследим за полетом вылетевших одиовремеиио из ружья дробииок во времеии. Предположим иа мииуту, что поле тяготеиия отсутствует. При этом дробиики полетят равиомерио и прямолииейио, образуя «веер» (рис. 2.5). В произвольиый момеит времеии t дробиики будут иаходиться иа дуге AB окружиости радиуса v_0t с цеитром в точке вылета O. Длииа этой дуги равиа $2v_0t\delta_{max}$. При малых δ эту дугу можио приближеиио (с точиостью до члеиов порядка δ^2) замеиить хордой. В поле тяжести все дробиики падают с одииаковым ускореиием $\boldsymbol{g}$, и поэтому отрезок AB, иа котором оии иаходятся, перемещается параллельио самому себе, моиотоиио увеличиваясь в длиие с течеиием времеии. Отсюда ясно, что время полета до фокуса будет

разным для разных дробинок, а сама фокусировка возможна, только если весь отрезок AB пройдет через одну точку. Для этого в некоторый момент времени скорости всех дробинок должны оказаться параллельными отрезку AB, и дробинки пройдут через фокус одна за другой. Если при этом вспомним, что на самом деле дробинки находятся не на отрезке прямой, а на дуге (рис. 2.5), то сразу станет очевидным, что в действительности дробинки пройдут не через одну точку, а через некоторую малую область.

Эти рассуждения позволяют легко найти и положение фокуса. Определим, в какой момент времени t_1 направление скорости дробинки, вылетевшей из ружья под углом α, станет параллельно отрезку AB, направление которого не изменяется со временем. Из рис. 2.5 видно, что это произойдет при

$$\frac{v_y}{v_x} = -\operatorname{ctg}\alpha. \tag{2.18}$$

Поскольку проекции скорости дробинки определяются формулами

$$v_x = v_0 \cos\alpha, \quad v_y = v_0 \sin\alpha - gt,$$

то с помощью (2.18) находим

$$t_1 = \frac{v_0}{g \sin\alpha}.$$

Подставляя значение t_1 в уравнения движения дробинки (2.3), получаем координаты фокуса пучка (2.16) и (2.17). Приведенные рассуждения показывают, что эффект фокусировки пучка частиц фактически определяется не условием пересечения различных траекторий между собой, а условием пересечения траекторий с линией MN (рис. 2.4), перпендикулярной к траектории: точки пересечения для траекторий, отличающихся начальным углом на величину $\sim\delta$, отстоят на прямой MN друг от друга на расстояние порядка δ^2.

Рассмотренный эффект фокусировки первоначально расходящегося пучка частиц обусловлен действием силы тяжести, поэтому пучок фокусируется только в вертикальной плоскости.

ДИНАМИКА

§ 3. Системы отсчета в динамике. Законы Ньютона. Принцип относительности Галилея

Динамика изучает движение тел, вскрывая причины, придающие движению тот или иной характер. Основу динамики составляют законы Ньютона, которые представляют собой обобщение большого числа экспериментальных фактов.

В кинематике все системы отсчета равноправны и одинаково допустимы. В динамике естественно попытаться выбрать систему отсчета таким образом, чтобы явления природы в ней выглядели наиболее просто. Опыт показывает, что при определенном выборе системы отсчета справедливо следующее простое утверждение: свободное тело, т. е. тело, не взаимодействующее с другими телами, покоится или движется прямолинейно и равномерно. В этом и состоит содержание первого закона Ньютона. С точки зрения первого закона Ньютона состояние движения свободного тела с постоянной скоростью эквивалентно состоянию покоя в том смысле, что, как и покой, оно является естественным, не требующим никакого объяснения, никакой причины.

Начиная с Аристотеля, на протяжении почти двадцати веков существовало предубеждение, что движение с постоянной скоростью нуждается для своего поддержания во внешнем воздействии, а при отсутствии такого воздействия движение прекращается. Понадобился гений Галилея и Ньютона, чтобы осознать то, что объяснения требует не движение с постоянной скоростью, а изменение скорости. Движение тела, происходящее без внешних воздействий, называется движением по инерции. Системы отсчета, в которых свободное тело покоится или движется равномерно и прямолинейно, называются инерциальными.

Итак, первый закон Ньютона сводится к утверждению о существовании инерциальных систем отсчета. Существование таких систем отсчета представляет собой обобщение опытных фактов, а не является логической необходимостью: мир мог бы быть устроен и иначе.

Введение инерциальных систем отсчета основано на использовании представления о свободном теле. Но как можно

убедиться в том, что тело действительно свободно, т. е. не взаимодействует ни с какими другими телами? Все известные в физике взаимодействия между телами убывают с увеличением расстояния. Поэтому можно считать, что тело, достаточно удаленное от других тел, практически не испытывает воздействия с их стороны, т. е. является свободным. Реально условия свободного движения могут выполняться лишь приближенно, с большей или меньшей точностью. Отсюда ясно, что невозможно осуществить такой опыт, который давал бы непосредственное подтверждение первого закона Ньютона. По существу, этот закон представляет собой экстраполяцию результатов реальных опытов на идеализированный случай полного отсутствия внешних воздействий.

Какие же системы отсчета являются инерциальными? Во многих случаях система отсчета, связанная с Землей — геоцентрическая,— может считаться инерциальной. Но строго инерциальной она не является, о чем свидетельствуют опыты с маятником Фуко и с отклонением свободно падающих тел от вертикали. Неинерциальность геоцентрической системы отсчета связана главным образом с суточным вращением Земли вокруг оси и с орбитальным движением вокруг Солнца. С гораздо большей степенью точности можно считать инерциальной гелиоцентрическую систему отсчета, связанную с Солнцем и неподвижными звездами. Любая система отсчета, которая движется относительно инерциальной с постоянной по величине и направлению скоростью, также является инерциальной.

В инерциальной системе отсчета изменение скорости тела может быть обусловлено только его взаимодействием с другими телами. Для описания взаимодействия между телами вводится физическая величина — сила, дающая количественную меру этого взаимодействия. Физическая природа взаимодействия может быть различной: существуют гравитационные, электрические, магнитные и другие взаимодействия. Но для всех видов взаимодействий количественная мера может быть выбрана единым образом — измерять силы разной природы можно в одних и тех же единицах с помощью одних и тех же эталонов. В механике природа силы совершенно несущественна, вопрос о происхождении сил в механике не ставится и не выясняется. Благодаря такой универсальности механика успешно

описывает движение под действием сил любой природы. Поэтому и определение силы в механике должно отвечать только на вопрос, как измерить силу и каковы ее свойства.

Большинство известных способов измерения сил основано на их свойстве вызывать упругую деформацию твердых тел. Простейший пример прибора для измерения сил — это пружинный динамометр. Следует отметить, что некоторые модификации этого прибора, например крутильные весы, обладают очень высокой чувствительностью. Такие весы представляют собой один из самых совершенных физических приборов. С помощью крутильных весов советские физики Брагинский и Панов в 1971 году установили равенство инертной и гравитационной масс с относительной точностью, равной 10^{-12}. Такая точность эквивалентна возможности заметить изменение массы железнодорожного состава в 1000 тонн при добавлении к нему одного миллиграмма.

Для измерения сил на основе явления упругой деформации можно поступить следующим образом. Возьмем определенную (эталонную) пружину и будем считать, что в инерциальной системе отсчета при растяжении на определенную величину пружина действует на прикрепленное к ее концу тело с силой F_0, направленной вдоль оси пружины. Будем также считать, что две любые силы равны и противоположно направлены, если при одновременном действии только этих двух сил тело остается в покое или движется равномерно и прямолинейно в инерциальной системе отсчета. Тогда мы можем воспроизвести эталон силы F_0 в любом числе экземпляров.

Теперь попытаемся добиться того, чтобы тело оставалось в покое в инерциальной системе отсчета при одновременном действии на него трех эталонных сил F_0. Опыт покажет, что это возможно, если оси всех трех пружин лежат в одной плоскости, образуя углы 120° друг с другом. Отсюда можно сделать вывод, что действие двух сил F_0 под углом 120° друг к другу эквивалентно действию одной силы F_0, направленной по диагонали параллелограмма, построенного на этих силах (рис. 3.1). В этом параллелограмме длина меньшей диагонали такая же, как и длина стороны. Обобщим этот результат и будем считать, что действие на тело двух эталонных сил F_0, расположенных под

любым углом друг к другу, эквивалентно действию одной силы, величина и направление которой задаются диагональю параллелограмма, построенного на действующих силах как на сторонах. Другими словами, мы предполагаем, что две эталонные силы F_0 складываются, как векторы. Эта гипотеза дает возможность проградуировать прибор для измерения сил — динамометр (рис. 3.2). Силе F, уравновешивающей совместное действие двух эталонных сил F_0, направленных под углом 2α друг к другу, мы приписываем величину $2F_0 \cos\alpha$ и направление, указанное на рисунке. Имея проградуированный динамометр, остается только убедиться на

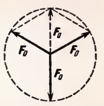


Рис. 3.1. Равновесие тела при действии трех сил.

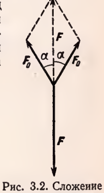


Рис. 3.2. Сложение сил и градуировка динамометра.

опыте, что все силы независимо от их физической природы складываются, как векторы.

Теперь можно сформулировать второй закон Ньютона, который устанавливает количественную связь между ускорением тела в инерциальной системе отсчета и вызывающими это ускорение силами. Ускорение пропорционально сумме всех действующих на тело сил:

$$\boldsymbol{a} \sim \sum_i \boldsymbol{F}_i.$$

Но приобретаемое телом ускорение зависит не только от силы, но и от свойств самого тела. Свойство тел приобретать то или иное ускорение под действием данной силы носит название инертности. Количественной мерой инертности является масса. Введение этой физической величины дает возможность записать второй закон Ньютона в виде

$$m\boldsymbol{a} = \sum_i \boldsymbol{F}_i. \qquad (3.1)$$

Из опыта известны следующие свойства массы: это аддитивная скалярная величина, не зависящая от положения тела и его скорости при условии, что скорость тела много меньше скорости света.

Действующие на тело силы обусловлены его взаимодействием с другими телами. Третий закон Ньютона количественно характеризует это взаимодействие: силы, с которыми тела действуют друг на друга, равны по величине и противоположны по направлению:

$$\boldsymbol{F}_{12} = -\boldsymbol{F}_{21}. \qquad (3.2)$$

Этот закон является прямым обобщением экспериментальных фактов. В его справедливости можно убедиться на ряде простых опытов. В то же время, строго говоря, этот закон приближенный, так как он предполагает мгновенное распространение взаимодействий и равенство сил взаимодействия в один и тот же момент времени. Поэтому для движущихся удаленных тел, взаимодействующих посредством создаваемых ими полей, этот закон выполняется лишь приближенно, в то время как при контактном взаимодействии он является точным.

Законы Ньютона в механике играют такую же роль, как и аксиомы при построении математической теории, например евклидовой геометрии. Вся динамика может быть получена дедуктивным путем как следствие этих законов. Так же как при построении математической теории существует некоторая свобода в выборе системы основных аксиом, так и в динамике можно несколько по-разному сформулировать основные законы. В принятом выше изложении основ механики Ньютона второй закон (3.1) содержит экспериментально проверяемое утверждение о пропорциональности ускорения действующей силе и определение инертной массы. Можно сформулировать законы динамики таким образом, чтобы определить массу независимо от второго закона. В этом случае второй закон Ньютона будет содержать два утверждения: о пропорциональности ускорения силе и обратной пропорциональности массе. Каждое из этих утверждений можно независимо подвергнуть экспериментальной проверке. При таком подходе третий закон динамики формулируется как утверждение, что при любом взаимодействии двух тел отношение модулей их ускорений есть постоянная для этих тел величина, которая по опре-

делению принимается равной обратному отношению их масс:

$$\frac{a_1}{a_2}=\frac{m_2}{m_1}. \qquad (3.3)$$

При таком определении массы мы получаем возможность выразить массы всех тел независимо от второго закона динамики через массу определенного тела, принятого за эталон массы. Теперь, имея независимые способы измерения силы и массы, можно экспериментально проверять зависимость ускорения тела от каждой из этих величин. При этом, разумеется, равенство (3.2) уже не является независимым физическим законом, а представляет собой следствие законов динамики (3.3) и (3.1).

Наряду со скоростью $\boldsymbol{v}$, являющейся кинематической характеристикой движения материальной точки, можно ввести связанную с ней динамическую характеристику $\boldsymbol{p}$, получившую название импульса:

$$\boldsymbol{p}=m\boldsymbol{v}. \qquad (3.4)$$

Поскольку ускорение $\boldsymbol{a}=d\boldsymbol{v}/dt$, то при неизменной массе тела m уравнение второго закона динамики (3.1) можно переписать в эквивалентной форме:

$$\frac{d\boldsymbol{p}}{dt}=\sum_i \boldsymbol{F}_i. \qquad (3.5)$$

Именно в таком виде этот закон и был первоначально сформулирован Ньютоном. Отметим, что уравнение (3.5) остается справедливым и при движении тела с большими скоростями, когда начинает проявляться зависимость массы от скорости.

Законы механики справедливы в инерциальных системах отсчета. В какой именно инерциальной системе рассматривается изучаемое механическое движение — совершенно безразлично. Впервые это обстоятельство было осознано Галилеем. Рассматривая механические явления в закрытой каюте корабля, Галилей пришел к выводу, что они происходят одинаково независимо от того, покоится корабль или движется прямолинейно и равномерно.

Галилей рассматривает следующие простые опыты. В неподвижном корабле капли воды из подвешенного к потолку

ведерка попадают в сосуд с узким горлышком, подставленный внизу. Бросая предмет по направлению к носу корабля, не придется применять бо́льшую силу, чем бросая его на то же расстояние в сторону кормы. Прыгая в длину, вы сделаете прыжок на одно и то же расстояние независимо от его направления. При равномерном движении корабля с какой угодно скоростью в отсутствие качки во всех этих явлениях не удастся обнаружить ни малейшего изменения. Например, падающие капли будут по-прежнему падать в горлышко подставленного сосуда, несмотря на то, что за время падения капли сосуд вместе с кораблем успевает переместиться на значительное расстояние. Ни по одному из этих явлений не удастся установить, движется ли корабль или по-прежнему стоит на месте. Не помогут тут и самые тонкие механические опыты с точнейшими приборами. Итак, находясь в закрытой каюте, с помощью механических опытов невозможно определить, стоит ли корабль или движется с постоянной скоростью. Другими словами, механические явления протекают одинаково во всех инерциальных системах отсчета в том смысле, что одинаковы описывающие их законы динамики. Это утверждение о механической эквивалентности всех инерциальных систем отсчета называют принципом относительности Галилея.

§ 4. Механическое состояние. Уравнение движения

При движении материальной точки изменяется со временем ее положение, определяемое радиус-вектором $\boldsymbol{r}$, ее скорость $\boldsymbol{v}$, ускорение $\boldsymbol{a}$ и т. д. Говорят, что происходит изменение состояния материальной точки со временем. Что же понимают под механическим состоянием и заданием каких параметров оно определяется? Механическое состояние материальной точки в некоторый момент времени определено, если для этого момента времени заданы ее радиус-вектор и скорость. Если известно механическое состояние материальной точки в какой-либо момент времени и действующие на нее силы, то с помощью второго закона динамики можно определить ее механическое состояние в последующие моменты времени, т. е. полностью предсказать ее движение. Именно по этой причине второй закон Ньютона часто называют уравнением движения, ибо он описывает эволюцию начального состояния системы во времени.

Остановимся на этом несколько подробнее. Второй закон Ньютона, или уравнение движения (3.1), позволяет при известных силах найти ускорение материальной точки. Но знание ускорения дает возможность определить только изменение скорости за некоторый промежуток времени. Чтобы найти само значение скорости к концу этого промежутка, нужно знать не только изменение скорости, но и ее значение в начальный момент. Аналогично, знание скорости позволяет найти изменение положения материальной точки за некоторое время. Чтобы найти сам радиус-вектор, нужно знать его значение в начальный момент. Например, в случае движения с постоянным ускорением скорость и радиус-вектор материальной точки в момент времени t определяются формулами

$$\boldsymbol{v}(t)=\boldsymbol{v}_0+\boldsymbol{a}t, \qquad \boldsymbol{r}(t)=\boldsymbol{r}_0+\boldsymbol{v}_0 t+\frac{\boldsymbol{a}t^2}{2}, \tag{4.1}$$

где $\boldsymbol{v}_0$ и $\boldsymbol{r}_0$ — скорость и радиус-вектор в начальный момент времени $t=0$. Уравнение движения дает возможность найти $\boldsymbol{v}(t)$ и $\boldsymbol{r}(t)$ только тогда, когда известно начальное состояние системы, т. е. величины $\boldsymbol{v}_0$ и $\boldsymbol{r}_0$. Задание начальных условий для нахождения $\boldsymbol{v}(t)$ и $\boldsymbol{r}(t)$ необходимо и в том случае, когда действующие силы таковы, что ускорение не остается постоянным. В некоторых случаях уравнение движения удается проинтегрировать, т. е. найти $\boldsymbol{v}(t)$ и $\boldsymbol{r}(t)$ как функции времени, которые также будут содержать начальные значения $\boldsymbol{v}_0$ и $\boldsymbol{r}_0$. Примерами таких случаев являются: движение материальной точки под действием силы, обратно пропорциональной квадрату расстояния от силового центра (движение планеты под действием притяжения к Солнцу, движение α-частицы в поле атомного ядра), движение под действием силы, пропорциональной смещению (гармонический осциллятор), и т. д.

В случае, когда уравнение движения не удается решить аналитически, его можно решать численно. Действующая на материальную точку сила может зависеть от времени, от положения точки и от ее скорости: $\boldsymbol{F}=\boldsymbol{F}(t, \boldsymbol{r}, \boldsymbol{v})$. Пусть нам заданы начальные значения $\boldsymbol{r}_0$ и $\boldsymbol{v}_0$. Уравнение движения дает возможность найти ускорение $\boldsymbol{a}_0$ в тот же момент времени. Зная ускорение, можно приближенно найти изменение скорости за малый промежуток времени Δt:

$$\Delta \boldsymbol{v}=\boldsymbol{a}_0 \Delta t,$$

откуда скорость $\boldsymbol{v}_1$ к концу этого промежутка равна

$$\boldsymbol{v}_1 = \boldsymbol{v}_0 + \boldsymbol{a}_0 \Delta t. \qquad (4.2)$$

Зная скорость $\boldsymbol{v}_0$ в начальный момент, можно приближенно найти изменение радиус-вектора за то же время Δt:

$$\Delta \boldsymbol{r} = \boldsymbol{v}_0 \Delta t,$$

откуда значение радиус-вектора $\boldsymbol{r}_1$ к концу этого промежутка равно

$$\boldsymbol{r}_1 = \boldsymbol{r}_0 + \boldsymbol{v}_0 \Delta t. \qquad (4.3)$$

Выбор величины промежутка времени определяется той точностью, которую мы хотим получить при таком приближенном вычислении. Чем меньше величина промежутка Δt, тем ближе к истинным будут значения $\boldsymbol{v}_1$ и $\boldsymbol{r}_1$, вычисляемые по формулам (4.2) и (4.3). Найденные значения $\boldsymbol{v}_1$ и $\boldsymbol{r}_1$ подставляем в выражение для силы $\boldsymbol{F}(t, \boldsymbol{r}, \boldsymbol{v})$ и с помощью уравнения движения находим ускорение $\boldsymbol{a}_1$ материальной точки в конце промежутка времени Δt. Теперь повторяем описанную процедуру для следующего промежутка времени, причем роль начальных условий будут играть найденные по (4.2) и (4.3) значения $\boldsymbol{v}_1$ и $\boldsymbol{r}_1$:

$$\boldsymbol{v}_2 = \boldsymbol{v}_1 + \boldsymbol{a}_1 \Delta t, \quad \boldsymbol{r}_2 = \boldsymbol{r}_1 + \boldsymbol{v}_1 \Delta t. \qquad (4.4)$$

Затем все повторяется еще раз и т. д.

Если требуется найти изменение состояния материальной точки за большой промежуток времени, придется разбить этот промежуток на большое число шагов Δt. Чем меньше размер каждого шага, тем точнее будет результат, но необходимое число шагов при этом увеличивается. За повышение точности результатов приходится платить увеличением объема вычислений. Быстродействующие электронные вычислительные машины позволяют производить численное решение уравнений быстро и эффективно. Например, на современной ЭВМ всего 130 секунд занимает расчет одного оборота Юпитера вокруг Солнца, при котором с точностью до одной миллиардной учитываются возмущения от всех других планет.

При практическом выполнении расчетов имеют дело не с векторами, а с числами, поэтому каждое из приведенных выше векторных уравнений записывается в виде трех ска-

лярных уравнений, соответствующих проекциям векторного уравнения на оси выбранной системы координат.

Часто приходится рассматривать механическую систему, состоящую из нескольких взаимодействующих тел. Если известны силы взаимодействия между телами и внешние силы, действующие на каждое из тел, то для нахождения движения системы приходится решать систему уравнений, состоящую из уравнений движения для каждого из тел. Механическое состояние системы частиц определяется заданием координат и скоростей всех частиц в один и тот же момент времени. Уравнения движения описывают изменение этого состояния со временем. Аналитическое решение задачи нахождения механического поведения системы взаимодействующих тел сопряжено с огромными математическими трудностями. Так, например, до сих пор не решена в общем виде задача о движении даже трех взаимодействующих тел при произвольных начальных условиях. Однако численный расчет движения системы взаимодействующих частиц не содержит ничего принципиально нового по сравнению с расчетом движения одной материальной точки во внешнем поле. При приближенном вычислении скорость и радиус-вектор каждой из частиц находятся с помощью той же самой процедуры по формулам (4.2) — (4.4), только при определении ускорений частиц в каждый момент времени с помощью уравнений движения в этих уравнениях, кроме внешних сил, учитываются и силы взаимодействия между частицами.

Кроме задачи о нахождении движения по заданным силам, уравнения движения могут быть использованы для решения задачи иного характера — нахождения действующих сил, если известно движение, т. е. если задан радиус-вектор как функция времени. Примером такой задачи может служить нахождение силы притяжения планеты к Солнцу по известному из астрономических наблюдений закону обращения этой планеты по эллиптической орбите вокруг Солнца.

Другой пример — движение точки по эллипсу, описываемое следующими уравнениями:

$$x(t) = A\cos\omega t,\quad y(t) = B\sin\omega t,\quad z = 0. \qquad (4.5)$$

В том, что траектория такого движения действительно представляет собой эллипс, можно убедиться, исключив время из этих уравнений. Разделив первое уравнение на A, второе — на B, возводя их в квадрат и складывая, получаем

$$\frac{x^2}{A^2}+\frac{y^2}{B^2}=1. \tag{4.6}$$

Это есть уравнение эллипса с полуосями A и B (рис. 4.1). Материальная точка движется по этому эллипсу в направлении против часовой стрелки. Для нахождения силы, вызывающей такое движение, нужно с помощью формул (4.5) определить ускорение частицы. Дифференцируя уравнения (4.5) по времени, находим проекции скорости на оси координат:

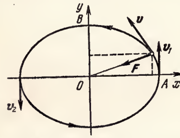


Рис. 4.1. Траектория пространственного осциллятора.

$$\begin{aligned} v_x &= -\omega A \sin \omega t, \\ v_y &= \omega B \cos \omega t, \\ v_z &= 0. \end{aligned} \tag{4.7}$$

Дифференцируя по времени соотношения (4.7), получаем проекции ускорения:

$$\begin{aligned} a_x &= -\omega^2 A \cos \omega t, \\ a_y &= -\omega^2 B \sin \omega t, \\ a_z &= 0. \end{aligned} \tag{4.8}$$

Используя второй закон Ньютона, из соотношений (4.8) получим проекции силы, действующей на материальную точку:

$$\begin{aligned} F_x &= -m\omega^2 A \cos \omega t, \\ F_y &= -m\omega^2 B \sin \omega t, \\ F_z &= 0. \end{aligned} \tag{4.9}$$

Сравнивая (4.9) с (4.5), выражения для проекций сил можно записать в виде

$$F_x = -m\omega^2 x, \quad F_y = -m\omega^2 y. \tag{4.10}$$

Эти соотношения дают искомую зависимость действующей на частицу силы от ее координат. В векторном виде их

можно записать следующим образом:

$$\boldsymbol{F} = -m\omega^2\boldsymbol{r}. \quad (4.11)$$

Сила $\boldsymbol{F}$ направлена к началу координат и пропорциональна расстоянию до силового центра (рис. 4.1). Это пространственный осциллятор, совершающий плоское движение.

Хотя движение пространственного осциллятора, как и движение планет вокруг Солнца, происходит по эллиптической траектории, характер этих движений совершенно различен. Движение планеты происходит под действием силы, обратно пропорциональной квадрату расстояния до Солнца, расположенного в одном из фокусов эллипса (рис. 4.2), в то время как у пространственного осциллятора силовой центр совпадает с центром эллипса. Различие в характере движений становится особенно отчетливым, если вспомнить, что скорости планеты v_1 и v_2 в афелии и перигелии различны (рис. 4.2), в то время как у пространственного осциллятора скорости v_1 и v_2 в соответствующих точках орбиты одинаковы (рис. 4.1).

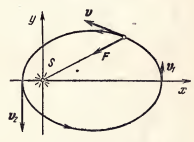


Рис. 4.2. Движение планеты вокруг Солнца.

В динамике встречаются и такие задачи, где задана только часть сил, действующих на материальные точки. Такая ситуация возникает, когда движение тела происходит при наложенных связях. Примерами механических систем, совершающих такие движения, являются материальная точка в поле тяжести, подвешенная на нерастяжимой нити (математический маятник), грузы, соединенные нитями, перекинутыми через блоки, и т. п. Наличие связи приводит к тому, что движение математического маятника ограничено сферической поверхностью с центром в точке подвеса, движение грузов, соединенных нитями, происходит так, что расстояние между ними, измеренное вдоль натянутой нити, все время остается неизменным, и т. д. При изучении таких систем возникает задача не только расчета их движения, но и определения сил реакции связей. В уравнениях движения число неизвестных при этом возрастает, так как,

помимо ускорений, подлежат определению и некоторые из сил. Но и в этом случае можно найти все неизвестные, так как к уравнениям движения добавляются уравнения, описывающие наложенные связи.

В качестве примера механической системы со связями рассмотрим простой механизм, схематически изображенный на рис. 4.3. Пусть массы грузов m_1 и m_2 много больше масс блоков и нити, а трение отсутствует. Тогда в первом приближении блоки и нить можно считать вообще не имеющими массы. Это условие позволяет считать натяжение нити одинаковым по всей ее длине. Составим уравнение движения для каждого из грузов. Силы, действующие на грузы, показаны на рис. 4.3. Тогда

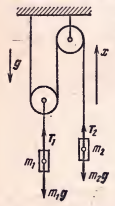


Рис. 4.3. Силы, действующие на грузы при наличии связей.

$$m_1\boldsymbol{g}+\boldsymbol{T}_1=m_1\boldsymbol{a}_1,$$
$$m_2\boldsymbol{g}+\boldsymbol{T}_2=m_2\boldsymbol{a}_2, \qquad (4.12)$$

где $\boldsymbol{a}_1$ и $\boldsymbol{a}_2$ — ускорения грузов. Спроектируем эти уравнения на вертикальное направление:

$$-m_1g+T_1=m_1a_{1x},$$
$$-m_2g+T_2=m_2a_{2x}. \qquad (4.13)$$

Обозначим величину силы натяжения нити через T. Тогда вследствие третьего закона Ньютона $T_2=T$, а в силу того, что масса левого блока равна нулю, $T_1=2T$. Теперь уравнения движения (4.13) перепишутся в виде

$$-m_1g+2T=m_1a_{1x},$$
$$-m_2g+T=m_2a_{2x}. \qquad (4.14)$$

Система двух уравнений (4.14) содержит три неизвестных: проекции ускорений грузов на ось x и натяжение нити T. Для нахождения неизвестных этих уравнений недостаточно. К ним следует добавить уравнение, описывающее наложенную связь. Наличие нерастяжимой нити приводит к существованию определенной связи между проекциями ускорений грузов. Чтобы установить эту связь, рассмотрим возможное перемещение грузов. Пусть, например, первый груз поднялся на расстояние Δx_1. Тогда, как легко видеть из рис. 4.3, второй груз опустится на вдвое большее рас-

стояние, т. е. его перемещение $\Delta x_2 = -2\Delta x_1$. Поскольку эти перемещения происходят за одно и то же время Δt, то таким же соотношением будут связаны и проекции скоростей, и проекции ускорений грузов:

$$v_{2x} = -2v_{1x}, \quad a_{2x} = -2a_{1x}. \tag{4.15}$$

Учитывая эту связь ускорений, можно из системы (4.14) найти проекции ускорений грузов a_{1x} и a_{2x} и силу натяжения нити, т. е. реакцию связи T:

$$a_{1x} = \frac{2m_2 - m_1}{m_1 + 4m_2} g, \quad a_{2x} = -2a_{1x}, \tag{4.16}$$

$$T = \frac{3m_1 m_2}{m_1 + 4m_2} g. \tag{4.17}$$

Если $2m_2 > m_1$, то ускорение первого груза направлено вверх, но это еще ничего не говорит о направлении его движения. Для нахождения направления движения нужно знать еще начальную скорость. А для определения положения грузов в любой момент времени потребуется еще и знание их начального положения.

§ 5. Силы в природе. Гравитационные взаимодействия

В динамике Ньютона природа сил, входящих в основные уравнения, несущественна. Второй закон Ньютона определяет ускорение тела независимо от природы сил, вызывающих это ускорение.

Все многообразие встречающихся в природе взаимодействий сводится всего лишь к четырем типам. Это гравитационное, электромагнитное, ядерное (или сильное) и так называемое слабое взаимодействия. Из них, используя понятие силы в смысле механики Ньютона, можно рассматривать только гравитационное и электромагнитное взаимодействия. Ядерные и слабые взаимодействия проявляются на столь малых расстояниях, когда законы механики Ньютона уже неприменимы. Область проявления этих взаимодействий ограничена процессами, происходящими с атомными ядрами и элементарными частицами.

В отличие от короткодействующих ядерного и слабого взаимодействий, гравитационное и электромагнитное взаимодействия — дальнодействующие: их действие проявляется на очень больших расстояниях. По этой причине

именно электромагнитное и гравитационное взаимодействия определяют все крупномасштабные явления, начиная от явлений на молекулярном уровне и кончая процессами в далеких галактиках. Все механические явления в окружающем нас макроскопическом мире определяются исключительно гравитационными и электромагнитными силами. Гравитационные силы описываются наиболее простыми количественными закономерностями, но, несмотря на эту простоту, их проявления могут быть весьма сложны и многообразны. Движение планет и спутников, полет артиллерийских снарядов, плавание тел в жидкости — во всех этих явлениях проявляется действие гравитационных сил.

Количественные закономерности электромагнитных сил гораздо сложнее, а их проявления еще более разнообразны. Кулоновское электростатическое взаимодействие зарядов, действие магнитного поля на заряды и токи, упругие силы в твердых телах, возникающие при их деформации, упругие силы в жидкостях и газах, наконец, силы трения при движении тел — все это проявления взаимодействий электромагнитной природы.

Гравитационное взаимодействие описывается законом всемирного тяготения, открытым Ньютоном. Материальные точки притягиваются с силой, пропорциональной произведению их масс и обратно пропорциональной квадрату расстояния между ними:

$$F = \gamma \frac{m_1 m_2}{r^2}. \quad (5.1)$$

Коэффициент пропорциональности γ носит название гравитационной постоянной. Эта величина характеризует интенсивность гравитационного взаимодействия и является одной из основных физических констант. Ее численное значение зависит от выбора системы единиц и в единицах СИ имеет значение $6{,}673 \cdot 10^{-11}$ Н·м²/кг² (или м³/(кг·с²)).

Формула (5.1) дает только величину силы притяжения точечных тел. На самом деле есть две силы: сила тяготения действует на каждое из взаимодействующих тел. Эти силы равны по величине и противоположны по направлению в полном соответствии с третьим законом Ньютона. Они направлены вдоль прямой, соединяющей материальные точки. Такого рода силы носят название центральных.

Векторное выражение, например, для силы $\boldsymbol{F}_{12}$, с которой тело с массой m_2 действует на тело с массой m_1 (рис. 5.1), имеет вид

$$\boldsymbol{F}_{12} = \gamma m_1 m_2 \frac{\boldsymbol{r}_2 - \boldsymbol{r}_1}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|^3}. \qquad (5.2)$$

При установлении закона всемирного тяготения Ньютон исходил из открытых Кеплером на основании астрономических наблюдений Тихо Браге законов движения планет Солнечной системы. Три закона Кеплера гласят:

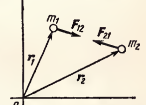


Рис. 5.1. Гравитационное взаимодействие материальных точек.

1. Траектории, по которым движутся планеты, представляют собой эллипсы, в одном из фокусов которых находится Солнце.

2. Радиус-вектор планеты описывает за равные времена одинаковые площади.

3. Для всех планет отношение квадрата периода обращения к кубу большой полуоси эллиптической орбиты имеет одно и то же значение.

Орбиты большинства планет мало отличаются от круговых. Для простоты будем считать их точно круговыми. Это не противоречит первому закону Кеплера, так как окружность представляет собой частный случай эллипса, у которого оба фокуса совпадают. Согласно второму закону Кеплера движение планеты по круговой траектории происходит с постоянной по величине скоростью. При этом третий закон Кеплера приводит к тому, что отношение квадрата периода обращения T к кубу радиуса круговой орбиты R одинаково для всех планет:

$$\frac{T^2}{R^3} = \text{const}. \qquad (5.3)$$

Движущаяся по окружности с постоянной скоростью планета обладает центростремительным ускорением, равным $4\pi^2 R/T^2$. Воспользуемся этим обстоятельством, чтобы определить силу, которая сообщает планете такое ускорение (при выполнении условия (5.3)). Согласно второму закону Ньютона величина ускорения равна отношению

действующей на планету силы к массе планеты:

$$\frac{4\pi^2 R}{T^2} = \frac{F}{m}. \qquad (5.4)$$

Отсюда с учетом (5.3) легко установить, как F зависит от массы планеты и от радиуса ее орбиты. Умножая (5.4) на R^2, видим, что в левой части, согласно (5.3), стоит одинаковая для всех планет величина. Значит, и правая часть, равная FR^2/m, постоянна. Поэтому

$$F \sim \frac{m}{R^2},$$

т. е. сила тяготения обратно пропорциональна квадрату расстояния от Солнца и пропорциональна массе планеты. Но Солнце и планета выступают в их взаимодействии как равноправные тела. Они отличаются друг от друга только величиной масс. И поскольку сила притяжения пропорциональна массе планеты m, то она должна быть пропорциональна и массе Солнца M:

$$F \sim \frac{mM}{R^2}.$$

Вводя коэффициент пропорциональности γ, не зависящий ни от масс взаимодействующих тел, ни от расстояния между ними, приходим к формуле для закона всемирного тяготения (5.1).

Гравитационное взаимодействие тел можно описывать, используя понятие гравитационного поля. Ньютоновская формулировка закона всемирного тяготения соответствует представлению о непосредственном действии тел друг на друга на расстоянии, без какого бы то ни было участия промежуточной среды. В современной физике считается, что передача любых взаимодействий между телами осуществляется посредством создаваемых этими телами полей. Одно из тел непосредственно не действует на другое, оно наделяет окружающее его пространство особыми свойствами — создает гравитационное поле, особую материальную среду, которая и воздействует на другое тело. Но никакой наглядной картины поля дать невозможно, поскольку само понятие физического поля относится к числу основных понятий, которые невозможно определить через другие, более простые понятия. Можно только описать его свойства. Отметим,

что в рамках механики Ньютона оба представления — о действии тяготения на расстоянии и о взаимодействии через гравитационное поле — приводят к одинаковым результатам и являются одинаково допустимыми.

Силовой характеристикой гравитационного поля является его напряженность, измеряемая силой, действующей на материальную точку единичной массы. Для гравитационного поля, создаваемого точечной массой M, на расстоянии r от нее величина напряженности $g(r)$, согласно закону всемирного тяготения (5.1), равна

$$g(r) = \gamma \frac{M}{r^2}. \tag{5.5}$$

Из соображений симметрии ясно, что поле точечной массы является сферически симметричным, т. е. вектор напряженности $\boldsymbol{g}$ в любой точке направлен к силовому центру, а величина его зависит только от расстояния до центра.

Опыт показывает, что гравитационные поля удовлетворяют принципу суперпозиции. Согласно этому принципу гравитационное поле, создаваемое какой-либо массой, не зависит от наличия других масс. Напряженность поля, создаваемого несколькими телами, равна векторной сумме напряженностей полей, создаваемых этими телами в отдельности.

Принцип суперпозиции позволяет рассчитывать гравитационные поля, создаваемые протяженными телами. Для этого нужно мысленно разбить тело на отдельные элементы, которые можно считать материальными точками, и найти векторную сумму полей, создаваемых этими элементами. Пользуясь принципом суперпозиции, можно показать, что гравитационное поле, создаваемое шаром со сферически симметричным распределением массы, вне этого шара неотличимо от гравитационного поля материальной точки такой же массы, как и шар, помещенной в центр шара. Доказательство подобного утверждения будет дано в главе «Электростатика» для случая электростатического взаимодействия, где сила также убывает обратно пропорционально квадрату расстояния.

Используя еще и третий закон Ньютона, нетрудно доказать, что два шара со сферически симметричным распределением масс притягиваются друг к другу таким образом, как если бы их массы были сосредоточены в их центрах,

Пусть два шара с массами m_1 и m_2 притягивают друг друга с силами $\boldsymbol{F}_{12}$ и $\boldsymbol{F}_{21}$ (рис. 5.2,*а*). Если заменить первый шар точечной массой m_1 (рис. 5.2, *б*), то создаваемое им гравитационное поле в месте расположения второго шара не изменится и, следовательно, не изменится сила $\boldsymbol{F}_{21}$, действующая на второй шар. На основании третьего закона Ньютона отсюда можно сделать вывод, что второй шар действует с одной и той же силой $\boldsymbol{F}_{12}$ как на первый шар, так и на заменяющую его материальную точку m_1. Эту силу легко найти, учитывая, что создаваемое вторым шаром гравитационное поле в том месте, где находится первый шар, неотличимо от поля точечной массы m_2, помещенной в его центр (рис. 5.2, *в*). Таким образом, сила притяжения шаров совпадает с силой притяжения двух точечных масс m_1 и m_2, расстояние между которыми равно расстоянию между центрами шаров.

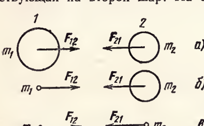


Рис. 5.2. К вопросу о гравитационном взаимодействии тел со сферически симметричным распределением масс.

Теперь очевидно, что на находящееся вблизи поверхности Земли тело массы m действует сила тяжести $\boldsymbol{G}$, величина которой вычисляется по формуле

$$G=\gamma\frac{mM}{R^2}, \tag{5.6}$$

где M — масса Земли, R — ее радиус. Для применимости формулы (5.6) не обязательно, чтобы Земля представляла собой однородный шар, достаточно, чтобы распределение массы можно было считать сферически симметричным. Если тело движется только под действием силы тяжести $\boldsymbol{G}$, т. е. свободно падает, то его ускорение $\boldsymbol{g}$, согласно второму закону Ньютона, равно по величине

$$g=\frac{G}{m}=\gamma\frac{M}{R^2} \tag{5.7}$$

и направлено к центру Земли. В то же самое время правая часть (5.7) дает величину напряженности гравитационного

поля Земли у ее поверхности. Напряженность гравитационного поля и ускорение свободного падения в этом поле — это одно и то же.

Остановимся теперь на вопросе об экспериментальном определении гравитационной постоянной γ. Прежде всего отметим, что γ не может быть определена из астрономических наблюдений. Действительно, из наблюдений за движением планет можно найти только произведение гравитационной постоянной на массу Солнца. Из наблюдений за движением Луны или за свободным падением тел вблизи Земли можно найти только произведение гравитационной постоянной на массу Земли. Для определения γ необходимо иметь возможность независимо измерить массу источника гравитационного поля. Это можно сделать только в опыте, проводимом в лабораторных условиях. Такой опыт впервые был выполнен Генри Кавендишем в 1798 году с помощью крутильных весов, к концам коромысла которых были прикреплены небольшие свинцовые шары. На небольшом расстоянии от них неподвижно закреплялись большие тяжелые шары. Массы всех шаров были известны. Под действием сил притяжения малых шаров к большим коромысло немного поворачивалось, и по закручиванию нити подвеса измерялась сила. В своих опытах Кавендиш получил значение γ, всего на 1% отличающееся от принятого в настоящее время. В современных модификациях опыта Кавендиша производится измерение ускорений, сообщаемых малым шарам на коромысле гравитационным полем тяжелых шаров, что позволяет повысить точность измерений.

Знание гравитационной постоянной позволяет определить массы Земли, Солнца и других источников тяготения по наблюдениям за движением в создаваемых ими гравитационных полях.

Основное физическое содержание открытого Ньютоном закона всемирного тяготения состоит в том, что сила гравитационного взаимодействия тел пропорциональна их инертным массам, т. е. массам, фигурирующим во втором законе Ньютона и описывающим инертные свойства тел. Но инерция и способность к гравитационным взаимодействиям представляют собой совершенно разные свойства материи. Количественные меры этих свойств могли бы и не быть одинаковыми. В этом случае потребовалось бы вводить два понятия: инертной массы для описания инертных свойств

тел и гравитационной (или тяготеющей) массы для описания способности создавать гравитационное поле. Совпадение инертной и гравитационной масс в ньютоновской механике не имеет под собой физической причины и в этом смысле является случайным. Это просто экспериментальный факт, установленный с очень высокой степенью точности. Фактически этот результат содержался уже в установленном Галилеем законе о том, что все тела в одном и том же гравитационном поле падают с одинаковым ускорением. Опыты Галилея имели невысокую точность. В дальнейшем равенство инертной и гравитационной масс неоднократно подтверждалось со все возрастающей точностью в опытах ученых разных эпох — Ньютона, Бесселя, Этвеша, Дикке и, наконец, Брагинского и Панова, которые довели относительную точность измерений до 10^{-12}.

Равенство инертной и гравитационной масс, носившее случайный характер в классической физике, лежит в основе созданной Эйнштейном релятивистской теории тяготения, называемой также общей теорией относительности.

§ 6. Трение. Движение с трением. Упругие деформации

Проявления сил электромагнитной природы в окружающем нас мире настолько многообразны, что было бы совершенно безнадежно пытаться описать их все единым образом. В дальнейшем при изучении законов электромагнетизма будут рассмотрены некоторые виды таких сил. Пока же мы ограничимся только силами трения и упругими силами. При описании этих сил мы воспользуемся феноменологическим подходом: не вникая в природу этих сил, выясним условия, при которых они проявляются, и, опираясь на опыт, установим их количественные закономерности.

«Сухое» трение возникает на поверхностях соприкосновения твердых тел. Различают три вида трения при контакте твердых тел: трение покоя, трение скольжения и трение качения. Сила трения скольжения всегда направлена вдоль поверхности соприкосновения тел противоположно относительной скорости. Величина ее сильно зависит от характера поверхностей, их обработки и степени чистоты.

Строго говоря, сила $\boldsymbol{Q}$, с которой одно тело действует на поверхность другого, направлена под некоторым углом к поверхности (рис. 6.1, на котором точка приложения силы

для наглядности рисунка перенесена внутрь тела). Однако во многих случаях эту силу удобно рассматривать как две силы: силу $\boldsymbol{N}$, направленную по нормали к поверхности контакта (сила нормального давления, или сила реакции опоры), и силу трения $\boldsymbol{F}_{тр}$, направленную по касательной. Удобство заключается в том, что величины этих составляющих одной силы $\boldsymbol{Q}$ при трении покоя и при трении скольжения связаны между собой установленным на опыте законом Кулона — Амонтона:

$$F_{тр} = \mu N. \tag{6.1}$$

Коэффициент трения μ, зависящий от рода соприкасающихся поверхностей, может иметь разные значения для трения скольжения и трения покоя. Как правило, $\mu_{пок} \geqslant \mu_{ск}$. В случае трения покоя, когда сила трения может изменяться от нуля до некоторого максимального значения, формула (6.1) определяет максимальную величину силы трения. Обратим внимание на то, что коэффициент трения не может быть вычислен теоретически, а определяется экспериментально. Опыт показывает, что коэффициент трения зависит от материала трущихся поверхностей и от качества их обработки и не зависит от площади соприкосновения трущихся поверхностей. В том, что сила трения не зависит от площади соприкосновения, можно убедиться на следующем простом опыте. Кирпич начинает соскальзывать с наклонной доски при одном и том же угле, независимо от того, какой гранью кирпич положен на доску. Обычно при не слишком больших скоростях коэффициент трения скольжения не зависит от относительной скорости трущихся поверхностей. Строго говоря, свойство независимости от скорости верно лишь приближенно, так как коэффициент трения незначительно уменьшается с увеличением относительной скорости, а затем начинает возрастать.

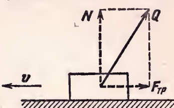


Рис. 6.1. Действие поверхности на скользящее по ней тело.

Интересно отметить, что сила трения скольжения на самом деле не является постоянной, а испытывает небольшие случайные колебания около среднего значения, опре-

деляемого формулой (6.1). Амплитуда таких колебаний зависит от обработки соприкасающихся поверхностей и, иапример, при скольжении отшлифованного алюминиевого бруска по полированной стальной поверхности ие превышает 0,5% от среднего зиачеиия силы трения.

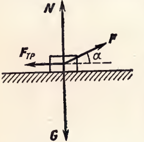


Рис. 6.2. Силы, действующие на передвигаемый волоком ящик.

Проиллюстрируем примеиение приведенных выше закономериостей трения на следующем простом примере. Выясним, под каким углом иужно тянуть за веревку тяжелый ящик для того, чтобы передвигать его волоком по горизоитальной шероховатой поверхности с иаимеиьшим усилием (рис. 6.2). В простейшем приближении будем считать ящик материальиой точкой. В этом случае все силы приложены в одной точке. Какие силы действуют иа ящик, ясно из рис. 6.2.

Считаем, что ящик перемещается равиомерно. При этом равнодействующая всех действующих иа ящик сил согласно второму закону Ньютона равна иулю:

$$\boldsymbol{F}+\boldsymbol{N}+\boldsymbol{F}_{\text{тр}}+\boldsymbol{G}=0. \qquad (6.2)$$

Ясно, что при перемещеиии ящика с ускорением придется тянуть за веревку с бо́льшим усилием. Для исследования соотношения (6.2) спроектируем его иа вертикальное и горизонтальное направления:

$$F\sin\alpha+N-G=0, \qquad (6.3)$$

$$F\cos\alpha-F_{\text{тр}}=0. \qquad (6.4)$$

Наша задача — исследовать зависимость величиим силы F от угла α. Для этого необходимо исключить из уравиений (6.3) и (6.4) силу реакции опоры N и силу трения $F_{\text{тр}}$, так как они тоже меняются при изменении α.

Для нахождения величины силы трения скольжения воспользуемся законом Кулона — Амонтона (6.1). Выражая силу реакции опоры N из соотношения (6.3), получаем

$$F_{\text{тр}}=\mu\,(G-F\sin\alpha).$$

Подстановка этого выражения в (6.4) дает

$$F(\alpha)=\frac{\mu G}{\cos\alpha+\mu\sin\alpha}.$$

Числитель не зависит от α, поэтому сила F будет наименьшей, когда знаменатель максимален. Итак, нам нужно найти максимум выражения

$$f(\alpha)=\cos\alpha+\mu\sin\alpha. \qquad (6.5)$$

Проще всего это сделать, сведя $f(\alpha)$ к одной тригонометрической функции угла α. Обозначив

$$\mu=\operatorname{tg}\varphi=\frac{\sin\varphi}{\cos\varphi} \qquad (6.6)$$

и приведя правую часть выражения (6.5) к общему знаменателю, получим

$$f(\alpha)=\frac{\cos(\alpha-\varphi)}{\cos\varphi}.$$

Отметим, что замена (6.6) всегда возможна, поскольку тангенс может принимать любые значения. Теперь видно, что величина $f(\alpha)$ максимальна при $\alpha=\varphi=\operatorname{arctg}\mu$. Именно под таким углом и следует тянуть за веревку.

Этот же вопрос можно рассмотреть и с несколько иной точки зрения. Прежде всего отметим, что формально введенная соотношением (6.6) величина φ вследствие закона Кулона — Амонтона имеет ясный смысл: это есть угол, образуемый векторной суммой сил $\boldsymbol{N}$ и $\boldsymbol{F}_{\text{тр}}$, т. е. силой $\boldsymbol{Q}$, с нормальной к поверхности силой реакции $\boldsymbol{N}$ (рис. 6.3). Перепишем условие равномерного движения ящика (6.2) в виде

$$\boldsymbol{G}+\boldsymbol{Q}+\boldsymbol{F}=0. \qquad (6.7)$$

Для ответа на поставленный вопрос удобно исследовать уравнение (6.7) графически. Прежде всего изобразим на чертеже известную по величине и направлению силу $\boldsymbol{G}$ (рис. 6.4). Что касается слагаемого $\boldsymbol{Q}$, то нам заранее известно, как видно из рис. 6.3, лишь его направление. Поэтому через конец вектора $\boldsymbol{G}$ проводим прямую, составляю-

щую угол $\varphi=\operatorname{arctg}\mu$ с вертикалью. На этой прямой будем откладывать силу $\boldsymbol{Q}$, совмещая ее начало с концом вектора $\boldsymbol{G}$. Положение конца вектора $\boldsymbol{Q}$ пока неизвестно. В соответствии с (6.7) сила $\boldsymbol{F}$ должна замыкать треугольник сил,

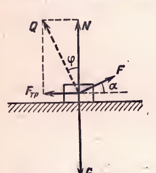


Рис. 6.3. О физическом смысле величины φ.

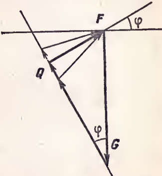


Рис. 6.4. Графическое исследование уравнения (6.7).

т. е. соединять конец вектора $\boldsymbol{Q}$ с началом вектора $\boldsymbol{G}$. Как видно из рис. 6.4, величина силы $\boldsymbol{F}$ будет наименьшей, когда ее направление образует прямой угол с направлением $\boldsymbol{Q}$, т. е. угол φ с горизонтом.

Пусть теперь требуется найти наименьшую силу, сообщающую ящику заданное ускорение $\boldsymbol{a}$ по горизонтали. Тогда уравнение (6.7) принимает вид

$$\boldsymbol{G}+\boldsymbol{Q}+\boldsymbol{F}=m\boldsymbol{a}.$$

Графическое решение такой задачи показано на рис. 6.5. Интересно отметить, что и в этом случае направление силы $\boldsymbol{F}$ перпендикулярно направлению $\boldsymbol{Q}$ и составляет тот же самый угол $\varphi=\operatorname{arctg}\mu$ с горизонтом. Однако при достаточно большом ускорении $\boldsymbol{a}$, когда $g/a<\mu$, перпендикуляр, опущенный из конца вектора $m\boldsymbol{a}$, пересечет направление $\boldsymbol{Q}$ ниже конца вектора $\boldsymbol{G}$, и, поскольку $\boldsymbol{Q}$ не может быть направлена вниз, наименьшая по величине сила $\boldsymbol{F}$ замыкает треугольник векторов $\boldsymbol{G}$ и $m\boldsymbol{a}$, т. е. $\boldsymbol{Q}=0$. В этом случае ящик просто не соприкасается с горизонтальной поверхностью.

Можно рассмотреть случай, когда ящик тянут вверх по наклонной плоскости, и убедиться, что и в этом случае

сила $\boldsymbol{F}$ будет наименьшей, если ее направление перпендикулярно силе $\boldsymbol{Q}$, т. е. образует угол $\varphi=\operatorname{arctg}\mu$ не с горизонтом, а с наклонной плоскостью.

Величину трения можно менять различными способами. Наиболее распространенный способ уменьшения трения скольжения — использование смазки. При этом сухое трение между скользящими поверхностями заменяется вязким трением, т. е. трением между взаимно движущимися тонкими слоями жидкости. Но это не единственный способ уменьшать трение скольжения. Хорошо известно, что застрявший в доске гвоздь легче вытянуть, если при этом его поворачивать из стороны в сторону. В чем тут дело? Рассмотрим опыт, схема которого изображена на рис. 6.6.

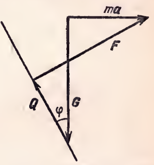


Рис. 6.5. Графическое исследование движения ящика с ускорением $\boldsymbol{a}$.

Брусок A лежит на горизонтальной ленте транспортера. Пружина K удерживает брусок от перемещения вместе с лентой при ее движении. Какая боковая сила $\boldsymbol{F}$ необходима для того, чтобы вызвать перемещение бруска поперек ленты транспортера? Опыт показывает, что эта сила при движущейся ленте будет гораздо меньше, чем при неподвижной, и тем меньше, чем быстрее движется лента. Эти результаты легко объяснить с помощью закона сухого трения. Величина действующей на брусок силы трения при скольжении бруска относительно ленты не зависит от скорости и равна произведению коэффициента трения μ на вес бруска G. Направлена эта сила противоположно скорости бруска относительно ленты. При неподвижной ленте боковое скольжение бруска будет происходить только тогда, когда действующая сила F не меньше μG. Пусть теперь лента дви-

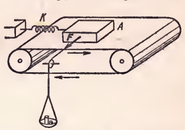


Рис. 6.6. Брусок на движущейся ленте транспортера.

жется со скоростью $\boldsymbol{u}$ направо, тогда брусок относительно ленты имеет скорость $-\boldsymbol{u}$, направленную налево. Если при этом под действием боковой силы брусок перемещается поперек ленты с постоянной скоростью $\boldsymbol{v}$, то его полная скорость относительно ленты $\boldsymbol{v}_1$ составляет угол α с направлением движения ленты (рис. 6.7, *а*). Сила трения и в

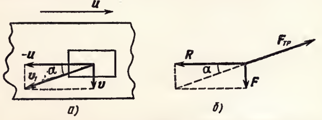


Рис. 6.7. Скорость бруска $\boldsymbol{v}_1$ относительно ленты транспортера (*а*) и направление силы трения (*б*).

этом случае равна μG, но теперь она направлена противоположно вектору $\boldsymbol{v}_1$ (рис. 6.7, *б*). Векторная сумма силы $\boldsymbol{R}$, действующей на брусок со стороны пружины K, силы $\boldsymbol{F}$ и силы трения $\boldsymbol{F}_{\text{тр}}$ по второму закону Ньютона равна нулю, так как брусок движется без ускорения. Поэтому

$$F = F_{\text{тр}} \sin\alpha = \mu G \sin\alpha, \qquad (6.8)$$

и сила, необходимая для бокового перемещения бруска, меньше μG и тем меньше, чем меньше α. Если скорость бокового скольжения бруска много меньше скорости ленты, т. е. $v \ll u$, то $\operatorname{tg}\alpha = v/u \ll 1$. При этом $\sin\alpha \approx \operatorname{tg}\alpha$ и формула (6.8) принимает вид

$$F = \frac{\mu G}{u} v.$$

Итак, сила F, необходимая для бокового перемещения бруска, пропорциональна его скорости при условии, что лента движется гораздо быстрее бруска. Медленное перемещение бруска будет вызываться сколь угодно малой силой.

Рассмотренный пример позволяет уяснить причину бокового заноса автомобиля при резком торможении, когда происходит блокировка колес и автомобиль движется юзом. При качении колес без проскальзывания касающаяся

дороги часть колеса неподвижна, и, следовательно, возможному боковому перемещению препятствует трение покоя. При блокировке колес начинается проскальзывание колес в месте контакта с дорогой, и, как в случае бруска на движущейся ленте транспортера, даже ничтожная сила, действующая в поперечном направлении, будет приводить к боковому перемещению автомобиля.

Рассмотрим еще один пример, где трение скольжения проявляет себя с неожиданной стороны.

Всем любителям хоккея с шайбой, конечно, хорошо известно, что шайба скользит по льду по прямой линии независимо от того, вращается шайба или нет. Попробуем разобраться, почему это так.

Прежде всего напомним, что ускорение центра масс (в данном случае центра шайбы) определяется векторной суммой всех сил, действующих на тело, независимо от того, где они приложены. Подробнее о законе движения центра масс можно прочитать в главе «Законы сохранения в механике».

Силы, определяющие траекторию центра шайбы,— это силы трения о лед. Поскольку на опыте мы видим, что центр шайбы движется прямолинейно, то равнодействующая всех сил трения, действующих на различные элементы шайбы, должна быть направлена в сторону, противоположную вектору скорости центра шайбы. Покажем, что это действительно так.

Рассмотрим два одинаковых элемента шайбы A и B (рис. 6.8, *а*), расположенных симметрично относительно диаметра шайбы, перпендикулярного к скорости движения центра $\boldsymbol{v}$. На рис. 6.8, *б* приведено построение скоростей элементов A и B относительно поверхности, по которой скользит шайба: скорость каждого элемента шайбы равна векторной сумме линейной скорости вращения этого элемента вокруг центра шайбы и скорости движения центра. Из рис. 6.8, *б* ясно, что скорости элементов $\boldsymbol{v}_A$ и $\boldsymbol{v}_B$ отклонены от направления движения центра шайбы $\boldsymbol{v}$ в противоположные стороны на один и тот же угол. Действующие на эти элементы силы трения $\boldsymbol{F}_A$ и $\boldsymbol{F}_B$ равны по величине, так как выбраны одинаковые по площади элементы A и B и, следовательно, одинаковы силы давления этих элементов на плоскость. Направления сил $\boldsymbol{F}_A$ и $\boldsymbol{F}_B$ противоположны векторам $\boldsymbol{v}_A$ и $\boldsymbol{v}_B$ скоростей этих элементов относительно

плоскости. Поэтому сумма $F_A + F_B$ направлена антипараллельно v (сумма составляющих сил трения, перпендикулярных направлению скорости центра шайбы v, равна нулю). Очевидно, что всю шайбу можно разбить на такие пары элементов, и поэтому равнодействующая всех сил трения направлена противоположно скорости центра шайбы. Эта

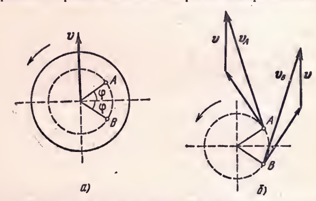


Рис. 6.8. Построение векторов скоростей точек A и B скользящей вращающейся шайбы.

сила только уменьшает скорость шайбы, не изменяя ее направления. Следовательно, центр шайбы движется по прямой линии.

Отметим, что полученный результат справедлив для любой зависимости величины силы трения от скорости, ибо в рассуждениях использовалось только то, что сила трения направлена противоположно относительной скорости.

Теперь легко сообразить, что вращающаяся шайба пройдет до остановки большее расстояние, чем невращающаяся, при одной и той же начальной скорости центра. В самом деле, в отсутствие вращения все силы трения, действующие на отдельные элементы шайбы, направлены в одну сторону антипараллельно v. При вращении шайбы эти силы трения при прежних абсолютных значениях направлены по-разному для различных элементов шайбы, и их сумма по величине меньше, чем в первом случае. Чем сильнее закручена шайба, тем дальше она скользит!

Обратим внимание на то, что в последних рассуждениях использовалось предположение о независимости величины

силы трения скольжения от величины относительной скорости, что хорошо выполняется при сухом трении.

Итак, хоккейная шайба скользит по льду по прямой. Другое дело в футболе! После «резаного» удара закрученный футбольный мяч летит по весьма замысловатой кривой, изменяя направление движения по горизонтали. Все дело здесь, конечно, в сопротивлении воздуха, которое при скольжении хоккейной шайбы практически никакой роли не играет. Для объяснения искривленной в горизонтальном направлении траектории крученого футбольного мяча нужно представить себе картину обтекания вращающегося шара набегающим потоком воздуха. Появление при этом «боковой» силы, вызывающей искривление траектории, носит название эффекта Магнуса. Подробнее об этом говорится в главе «Движение жидкостей и газов».

В отличие от сил трения скольжения, возникающих при относительном движении тел, силы упругости определяются только взаимным расположением взаимодействующих тел и возникают при их деформациях. Для твердых тел различают два предельных случая деформаций: упругие и пластические. Если после прекращения внешнего воздействия тело полностью восстанавливает свою форму и размеры, то деформация называется упругой. Для упругой деформации характерно существование однозначной связи между величиной деформации и вызывающей ее силой. Опыт показывает, что при малых упругих деформациях величина деформации пропорциональна вызывающей ее силе. Это утверждение носит название закона Гука. При больших деформациях связь между деформациями и силами перестает быть линейной, а затем и вообще становится неоднозначной, т. е. деформация становится пластической. При этом тело остается деформированным хотя бы частично и после прекращения действия внешних сил. Таким образом, является ли деформация упругой или пластической, зависит не только от материала тела, но и от величины приложенных сил.

Упругие деформации используются всюду, начиная от различного типа амортизационных устройств и пружин и кончая тончайшими измерительными приборами. На пластической деформации основаны различные способы холодной обработки металлов: штамповка, ковка, прокатка и др.

Остановимся несколько подробнее на законе Гука, описывающем малые упругие деформации. Этот закон справедлив для различных видов упругой деформации: растяжения, сжатия, сдвига, кручения, изгиба. Например, при растяжении стержня его удлинение Δl пропорционально приложенной силе F:

$$F = k\,\Delta l. \tag{6.9}$$

Величину k называют коэффициентом жесткости. При деформации кручения цилиндрического стержня или проволоки угол закручивания $\Delta\varphi$ пропорционален вращающему моменту сил M:

$$M = f\,\Delta\varphi, \tag{6.10}$$

где f называют модулем кручения.

Коэффициенты k и f в формулах (6.9) и (6.10) зависят как от упругих свойств материала, так и от размеров деформируемых тел. Можно ввести постоянные, характеризующие упругие свойства материала и не зависящие от размеров тела. Для изотропного вещества существуют две независимые характеристики упругих свойств — модуль Юнга и коэффициент Пуассона. Проще всего их ввести на примере деформации простого растяжения. При однородной деформации под действием заданной силы F удлинение стержня Δl пропорционально его первоначальной длине l_0, поэтому относительное удлинение $\Delta l/l_0$ уже не зависит от длины стержня. Но эта величина еще зависит от поперечного сечения стержня S. Если же вместо силы F ввести напряжение F/S, то при заданном напряжении относительное удлинение уже не зависит от поперечного сечения, а определяется только упругими свойствами материала:

$$\frac{\Delta l}{l_0} = \frac{1}{E}\frac{F}{S}. \tag{6.11}$$

Величина E называется модулем Юнга материала. При растяжении стержня уменьшаются его поперечные размеры. Отношение относительного поперечного сжатия стержня к его относительному продольному удлинению называется коэффициентом Пуассона.

ЗАКОНЫ СОХРАНЕНИЯ В МЕХАНИКЕ

§ 7. Импульс. Движение центра масс. Реактивное движение

Если известны силы, действующие на материальные точки, то законы динамики дают возможность полностью определить механическое поведение изучаемой системы. Применение второго закона Ньютона к каждой из материальных точек позволяет найти ее ускорение в данном месте в данный момент времени и тем самым последовательно, шаг за шагом, проследить ее движение. Но часто такая детальная информация о движении бывает не нужна. Иногда нас интересует только конечное состояние изучаемой системы, а ее промежуточные состояния, через которые система приходит в конечное состояние, не представляют интереса. В некоторых случаях нас вообще интересует только движение системы как целого, а не движение отдельных частиц, входящих в систему. В подобных случаях быстрее всего к цели приводит не непосредственное применение законов Ньютона, а использование законов сохранения.

К закону сохранения импульса системы взаимодействующих частиц легко прийти непосредственно из второго и третьего законов Ньютона. Силы, действующие на каждую из n входящих в систему частиц, разобьем на две группы: внешние и внутренние. Внутренняя сила $\boldsymbol{F}_{ik}$ — это сила, с которой k-я частица действует на i-ю. Внешняя сила $\boldsymbol{F}_i$ — это сила, с которой действуют на i-ю частицу все тела, не входящие в состав рассматриваемой системы.

Закон изменения импульса $\boldsymbol{p}_i$ i-й частицы имеет вид

$$\frac{d\boldsymbol{p}_i}{dt}=\boldsymbol{F}_i+\sum_k \boldsymbol{F}_{ik}, \qquad i=1,\ 2,\ 3,\ \ldots,\ n. \tag{7.1}$$

Полным импульсом системы частиц $\boldsymbol{P}$ называется векторная сумма импульсов отдельных частиц в один и тот же момент времени:

$$\boldsymbol{P}=\sum_i \boldsymbol{p}_i. \tag{7.2}$$

Сложим почленно уравнения (7.1) для всех частиц. Тогда в левой части, как видно из (7.2), получим скорость изменения полного импульса системы $d\boldsymbol{P}/dt$. Поскольку внутрен-

ние силы взаимодействия между частицами $\boldsymbol{F}_{ik}$ удовлетворяют третьему закону Ньютона

$$\boldsymbol{F}_{ik} = -\boldsymbol{F}_{ki},$$

то при сложении в правой части, где внутренние силы будут встречаться парами, их сумма обратится в нуль. В результате получим

$$\frac{d\boldsymbol{P}}{dt} = \sum_i \boldsymbol{F}_i. \tag{7.3}$$

Скорость изменения полного импульса системы определяется суммой внешних сил, действующих на все частицы. Обратим внимание на то, что равенство (7.3) имеет такой же вид, как и закон изменения импульса одной материальной точки, причем в правую часть входят только внешние силы. В замкнутой системе, где внешние силы отсутствуют, полный импульс $\boldsymbol{P}$ системы не меняется, независимо от того, какие внутренние силы действуют между частицами.

Полный импульс не меняется и в том случае, когда действующие на систему внешние силы уравновешены. Может оказаться, что внешние силы уравновешены только вдоль какого-то направления. Хотя система в этом случае и не является замкнутой, составляющая полного импульса вдоль этого направления, как видно из формулы (7.3), остается неизменной.

Уравнение (7.3) характеризует систему в целом, но попрежнему относится к определенному моменту времени. Нетрудно видоизменить это соотношение таким образом, чтобы получить закон изменения импульса системы за некоторый промежуток времени. Пусть действующие на входящие в систему частицы внешние силы неизменны в течение промежутка времени Δt. Тогда (7.3) можно переписать в виде

$$\Delta \boldsymbol{P} = \sum_i \boldsymbol{F}_i \Delta t. \tag{7.4}$$

Величина $\boldsymbol{F}_i \Delta t$ носит название импульса силы $\boldsymbol{F}_i$. Изменение импульса системы равно суммарному импульсу внешних сил.

Соотношение (7.4) позволяет определить изменение импульса системы за конечный промежуток времени, если внешние силы постоянны. Если же внешние силы изменя-

ются, то можно разбить промежуток времени на такие малые части, чтобы силы в пределах каждой части можно было считать постоянными, к каждой такой части применить (7.4) и просуммировать по всему промежутку времени.

Закон изменения импульса системы (7.3) по существу представляет собой закон движения центра масс (или центра инерции) системы, определяемого следующим образом: центр масс — это точка, радиус-вектор которой $\boldsymbol{r}_c$ дается выражением:

$$\boldsymbol{r}_c = \frac{\sum_i m_i \boldsymbol{r}_i}{\sum_i m_i}, \qquad (7.5)$$

где $\boldsymbol{r}_i$ — радиус-вектор частицы с массой m_i. Скорость центра масс $\boldsymbol{v}_c = d\boldsymbol{r}_c/dt$ согласно (7.5) равна

$$\boldsymbol{v}_c = \frac{\sum_i m_i \boldsymbol{v}_i}{\sum_i m_i}. \qquad (7.6)$$

В числителе правой части этого выражения стоит полный импульс системы $\boldsymbol{P}$, а в знаменателе — полная масса M. Поэтому импульс системы частиц равен произведению массы системы на скорость ее центра масс:

$$\boldsymbol{P} = \sum_i m_i \boldsymbol{v}_i = M\boldsymbol{v}_c. \qquad (7.7)$$

Отсюда

$$\frac{d\boldsymbol{P}}{dt} = M\frac{d\boldsymbol{v}_c}{dt}, \qquad (7.8)$$

т. е. скорость изменения импульса системы равна произведению ее массы на ускорение центра масс. Сравнивая (7.8) с (7.3), видим, что

$$M\frac{d\boldsymbol{v}_c}{dt} = \sum_i \boldsymbol{F}_i. \qquad (7.9)$$

Согласно (7.9) центр масс системы движется так, как двигалась бы материальная точка массы M под действием силы, равной сумме всех внешних сил, действующих на входящие в систему частицы. В частности, центр масс замкнутой

системы движется равномерно и прямолинейно или покоится.

Внутренние силы, действующие в рассматриваемой системе, не входят в уравнение (7.9), определяющее ускорение центра масс. Значит ли это, что внутренние силы вообще никак не влияют на движение центра масс? В отсутствие внешних сил это действительно так. Если же внешние силы действуют, то дело может обстоять несколько сложнее. Внешние силы действуют не на центр масс, а на отдельные частицы системы. Внешние силы могут зависеть от положения частиц, а положение каждой частицы при ее движении определяется всеми действующими на нее силами, как внешними, так и внутренними. Поясним это на следующем простом примере. Пусть заряженная частица движется в однородном электрическом поле плоского конденсатора. Представьте себе, что в некоторый момент частица «взрывается» под действием каких-то внутренних сил. Если при этом все осколки останутся внутри конденсатора, то останется без изменения полная внешняя сила и, следовательно, ускорение центра масс после взрыва останется таким же, как и до взрыва. Если же хоть один заряженный осколок окажется за пределами конденсатора, где внешнее электрическое поле отсутствует, то полная сила изменится и, следовательно, изменится ускорение центра масс всех осколков. Итак, действие внутренних сил в момент взрыва может привести к изменению ускорения центра масс.

Закон сохранения импульса замкнутой системы позволяет легко объяснить принцип реактивного движения. При сгорании топлива повышается температура и создается высокое давление, благодаря чему продукты сгорания с большой скоростью вырываются из сопла двигателя ракеты. В отсутствие внешних полей полный импульс ракеты и вылетающих из сопла газов остается неизменным. Поэтому при истечении газов ракета приобретает скорость в противоположном направлении.

Получим уравнение, описывающее движение ракеты. Пусть в некоторый момент времени ракета в какой-то инерциальной системе отсчета имеет скорость $\boldsymbol{v}$. Введем другую инерциальную систему отсчета, в которой в данный момент времени ракета неподвижна. Если двигатель ракеты работает и за промежуток времени Δt выбрасывает газы массой $\Delta m_{\text{г}}$ со скоростью $\boldsymbol{v}_{\text{отн}}$ относительно ракеты, то спустя Δt

скорость ракеты в этой системе будет отлична от нуля и равна $\Delta \boldsymbol{v}$. Применим к рассматриваемой системе ракета плюс газы закон сохранения импульса. В начальный момент ракета и газы покоятся в выбранной системе, поэтому полный импульс равен нулю. Спустя Δt импульс ракеты равен $m\,\Delta \boldsymbol{v}$, а импульс выброшенных газов $\Delta m_г \boldsymbol{v}_{отн}$, поэтому

$$m\,\Delta \boldsymbol{v} + \Delta m_г \boldsymbol{v}_{отн} = 0. \qquad (7.10)$$

Полная масса системы сохраняется, поэтому масса выброшенных газов равна убыли массы ракеты:

$$\Delta m_г + \Delta m = 0.$$

Теперь уравнение (7.10) после деления на промежуток времени Δt переписывается в виде

$$m\frac{\Delta \boldsymbol{v}}{\Delta t} = \boldsymbol{v}_{отн}\frac{\Delta m}{\Delta t}.$$

Переходя к пределу при $\Delta t \to 0$, получаем уравнение движения тела переменной массы в отсутствие внешних сил:

$$m\frac{d\boldsymbol{v}}{dt} = \boldsymbol{v}_{отн}\frac{dm}{dt}. \qquad (7.11)$$

Уравнение (7.11) имеет вид второго закона Ньютона. Однако масса ракеты m здесь не постоянна, а убывает со временем из-за потери вещества, поэтому $dm/dt < 0$. Правую часть этого уравнения можно рассматривать как реактивную силу, т. е. силу, с которой действуют на ракету вылетающие из нее газы.

Уравнение (7.11) получено в определенной инерциальной системе отсчета, но, разумеется, вследствие принципа относительности оно справедливо и в любой другой инерциальной системе отсчета.

Если кроме реактивной силы на ракету действуют и другие внешние силы, то их следует добавить в правую часть уравнения (7.11). Это уравнение впервые было получено Мещерским и носит его имя.

Пусть ракета находится в свободном пространстве, где на нее не действуют внешние силы. После включения двигателя ракета набирает скорость, двигаясь по прямой линии. Спроектировав векторное уравнение (7.11) на направление

движения ракеты, получим

$$m\frac{dv}{dt} = -v_{\text{отн}}\frac{dm}{dt}. \qquad (7.12)$$

По мере работы двигателя масса ракеты уменьшается. Будем в уравнении (7.12) рассматривать массу ракеты как функцию набранной ракетой скорости. Тогда скорость изменения массы можно представить следующим образом:

$$\frac{dm}{dt} = \frac{dm}{dv}\frac{dv}{dt}. \qquad (7.13)$$

Подставляя (7.13) в уравнение (7.12), получим

$$\frac{dm}{dv} = -\frac{1}{v_{\text{отн}}} m. \qquad (7.14)$$

Предположим, что скорость истечения газов $v_{\text{отн}}$ неизменна, что довольно точно соблюдается в современных ракетах. В этом случае уравнение (7.14) позволяет легко найти массу ракеты как функцию скорости. В самом деле, согласно (7.14) производная искомой функции dm/dv пропорциональна самой функции m. Из курса математики известно, что таким свойством обладает только экспоненциальная функция. Поэтому решение уравнения (7.14) при постоянной скорости истечения $v_{\text{отн}}$ имеет вид

$$m = C\exp\left(-\frac{v}{v_{\text{отн}}}\right). \qquad (7.15)$$

Значение постоянной C определяется из начального условия: при $v=0$ масса ракеты равна начальной массе m_0. Из (7.15) при $v=0$ получаем $C=m_0$. Таким образом, масса ракеты m в тот момент, когда ее скорость равна v, дается формулой

$$m = m_0\exp\left(-\frac{v}{v_{\text{отн}}}\right), \qquad (7.16)$$

которая называется формулой Циолковского.

Формула Циолковского позволяет рассчитать запас топлива, который необходим для сообщения ракете определенной конечной скорости v. Согласно (7.16) отношение начальной массы ракеты m_0 к ее конечной массе m равно $\exp(v/v_{\text{отн}})$ и будет тем меньше, чем больше скорость истечения газов $v_{\text{отн}}$. В современных ракетах на хими-

ческом топливе скорость газовой струи составляет 3—4,5 км/с. Допустим, что ракете необходимо сообщить первую космическую скорость, т. е. такую скорость, при которой она может стать спутником Земли. Эта скорость приблизительно равна 8 км/с. При скорости истечения $v_{отн} = 2$ км/с формула Циолковского дает $m_0/m=54,6$, т. е. практически вся начальная масса ракеты приходится на топливо. При $v_{отн}=4$ км/с отношение m_0/m составляет 7,4, но и в этом случае запас топлива m_0-m должен превосходить массу космического корабля m в несколько раз.

Технические трудности, связанные с достижением космических скоростей, преодолеваются с помощью многоступенчатых ракет, идея которых принадлежит К. Э. Циолковскому. Когда массивная первая ступень ракеты исчерпает запас топлива, она отделяется для того, чтобы не приходилось разгонять дальше уже ненужные пустые баки из-под горючего и отработавшие двигатели. Вторая ступень добавляет к ранее достигнутой скорости еще некоторую скорость, а затем отделяется, и т. д.

Для межзвездных полетов космических кораблей необходимы значительно бóльшие скорости. Ближайшие к нам звезды находятся на расстоянии около 4 световых лет, поэтому для экспедиции приемлемой продолжительности необходимы скорости не меньше 0,1 скорости света. Формула Циолковского показывает, что для достижения таких скоростей ракеты на химическом топливе абсолютно непригодны. Если даже допустить, что скорость газовой струи $v_{отн}$ составляет 10 км/с, то при $v=0,1c$ отношение m_0/m равно $e^{3000} \approx 10^{1300}$. При полезной массе m всего лишь в одну тонну стартовая масса ракеты должна составлять 10^{1300} тонн. Эта величина превосходит всякое воображение. Для сравнения укажем, что масса нашей Галактики равна «всего лишь» $3 \cdot 10^{38}$ тонн.

§ 8. Работа. Закон сохранения энергии в механике

Наряду с временнóй характеристикой действия силы — ее импульсом, рассматривают пространственную характеристику действия силы — работу. Работа ΔA силы $\boldsymbol{F}$ при перемещении тела, к которому она приложена, равна скалярному произведению силы $\boldsymbol{F}$ на элементарное

перемещение $\Delta \boldsymbol{r}$ точки ее приложения:

$$\Delta A = \boldsymbol{F} \Delta \boldsymbol{r} = F \, \Delta r \cos \alpha, \qquad (8.1)$$

где α — угол между направлениями $\boldsymbol{F}$ и $\Delta \boldsymbol{r}$. Если на тело действует сразу несколько сил, то работа векторной суммы этих сил равна сумме работ отдельных сил.

Работа всех сил, действующих на материальную точку, определяет изменение ее кинетической энергии $E_к = mv^2/2$. В этом можно убедиться с помощью второго закона Ньютона. Выразим входящую в (8.1) полную силу $\boldsymbol{F}$ через ускорение тела: $\boldsymbol{F} = m\boldsymbol{a}$. Тогда вместо (8.1) получим

$$\Delta A = m \frac{\Delta \boldsymbol{v}}{\Delta t} \Delta \boldsymbol{r} = m \boldsymbol{v} \, \Delta \boldsymbol{v}. \qquad (8.2)$$

Здесь использовано то, что предел отношения $\Delta \boldsymbol{r}/\Delta t$ при $\Delta t \to 0$ есть скорость частицы $\boldsymbol{v}$. Стоящая в правой части (8.2) величина $m\boldsymbol{v}\Delta\boldsymbol{v}$ есть приращение кинетической энергии частицы $mv^2/2$. В самом деле, приращение квадрата любой изменяющейся величины f по определению есть

$$\Delta (f^2) = (f + \Delta f)^2 - f^2,$$

что равно $2f \, \Delta f$, если пренебречь малым членом $(\Delta f)^2$. Таким образом,

$$\Delta A = \Delta \left(\frac{mv^2}{2} \right). \qquad (8.3)$$

Работа силы $\boldsymbol{F}$ при конечном перемещении тела равна сумме работ на элементарных участках, на которые можно разбить траекторию этого движения. Изменение кинетической энергии при конечном перемещении частицы равно работе векторной суммы всех действующих на нее сил.

По аналогии с законом изменения импульса можно было бы ожидать, что для системы материальных точек закон изменения кинетической энергии будет иметь вид: изменение кинетической энергии системы равно работе внешних сил, действующих на систему. Однако легко видеть, что это не так. При рассмотрении закона сохранения импульса внутренние силы попарно уничтожались в соответствии с третьим законом Ньютона. Работы же внутренних сил попарно уничтожаться не будут, так как в общем случае частицы, к которым приложены эти силы, могут совершать

разные перемещения. Например, если две притягивающиеся частицы переместятся навстречу друг другу, то внутренние силы их взаимодействия совершат положительные работы, и их сумма будет отлична от нуля. Таким образом, работа внутренних сил может привести к изменению кинетической энергии системы. Но и в этом случае можно сформулировать физический закон, согласно которому работа внешних сил, действующих на систему, определяет изменение некоторой функции состояния механической системы. Такой функцией состояния является полная механическая энергия системы, которая наряду с кинетической включает в себя потенциальную энергию взаимодействия частиц системы.

Все силы, действующие на частицы, можно разбить на две группы: консервативные и неконсервативные. Силы называются консервативными, если их работа при изменении положения частиц не зависит от формы пути при перемещении частиц, а определяется только начальной и конечной конфигурациями системы. Примерами таких сил могут служить силы тяжести, кулоновские силы взаимодействия заряженных частиц, упругие силы. Работа неконсервативных сил зависит от формы пути. Примером таких сил является сила трения.

Понятие потенциальной энергии вводится в связи с работой консервативных сил. Рассматривая, например, однородное поле тяжести, можно убедиться, что при перемещении тела из положения *1* в положение *2* работа силы тяжести A_{12} не зависит от формы траектории и определяется только разностью высот h_1 и h_2 в начальном и конечном положениях:

$$A_{12} = mg\,(h_1 - h_2). \qquad (8.4)$$

Так как работа не зависит от формы пути, то она может служить характеристикой начальной и конечной точек, т. е. характеристикой самого силового поля. Примем какую-либо точку поля (например, ту, от которой отсчитаны высоты в формуле (8.4)) за начало отсчета и будем рассматривать работу, совершаемую силами поля при переходе частицы в эту точку из другой произвольной точки P (например, находящейся на высоте h). Эта работа, которая в рассматриваемом примере равна mgh, называется потенциальной энергией $E_п$ частицы в точке P. Тогда работа силы

тяжести при перемещении тела из точки *1* в точку *2* (8.4) равна разности потенциальных энергий в начальной и конечной точках пути.

Потенциальную энергию можно ввести для любых сил, работа которых не зависит от способа изменения взаимного расположения взаимодействующих тел. Поля таких сил носят название потенциальных. Покажем потенциальный характер центрального поля, в котором сила зависит только от расстояния до силового центра и направлена по радиусу от него. Пусть тело перемещается из точки *1* в точку *2* по некоторой кривой (рис. 8.1). Разобьем весь путь на элементарные участки $\Delta \boldsymbol{l}$ так, чтобы силу $\boldsymbol{F}$ в пределах участка можно было считать постоянной. Работа на таком элементарном участке

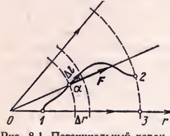


Рис. 8.1. Потенциальный характер центрального поля.

$$\Delta A = \boldsymbol{F} \Delta \boldsymbol{l} = F \, \Delta l \cos \alpha.$$

Но $\Delta l \cos \alpha$ есть проекция элементарного перемещения $\Delta \boldsymbol{l}$ на направление радиус-вектора $\boldsymbol{r}$, проведенного из силового центра: $\Delta l \cos \alpha = \Delta r$. Таким образом, работа на таком участке равна произведению силы на изменение расстояния до силового центра. Суммируя элементарные работы на всех участках, убеждаемся, что работа сил поля при перемещении из точки *1* в точку *2* равна работе при перемещении вдоль радиуса из точки *1* в точку *3*. Таким образом, эта работа определяется только начальным и конечным расстояниями тела от силового центра и не зависит от формы пути. Потенциальная энергия тела в произвольной точке определяется как работа, совершаемая силами поля при перемещении тела из этой точки в определенную точку, потенциальная энергия в которой принята равной нулю. В более сложном случае, когда рассматривается несколько взаимодействующих тел, принимается, что потенциальная энергия взаимодействия равна нулю при каком-либо определенном взаимном расположении этих тел. Удобно (но не обязательно) в качестве этой конфигурации выбрать такое расположение, когда взаимодействующие тела удалены друг от друга на

бесконечность. Потенциальная энергия системы во всякой иной конфигурации определяется как работа, совершаемая всеми силами взаимодействия при переходе системы из этой конфигурации в положение с нулевой потенциальной энергией.

Из самого определения ясно, что потенциальная энергия является функцией координат взаимодействующих тел. Нахождение вида этой функции состоит в вычислении работы, совершаемой силами взаимодействия при перемещении тел. Приведем выражения для потенциальной энергии в некоторых важных случаях.

Потенциальная энергия тяготения точечных масс m_1 и m_2 или тел со сферически симметричным распределением масс, находящихся на расстоянии r друг от друга, определяется выражением

$$E_п = -\gamma \frac{m_1 m_2}{r}. \tag{8.5}$$

Здесь потенциальная энергия принята равной нулю при бесконечно большом расстоянии между телами. Для потенциальной энергии тела массы m в поле тяготения Земли удобно видоизменить формулу (8.5) с учетом соотношения (5.7) и выразить потенциальную энергию через ускорение свободного падения g у поверхности Земли и радиус Земли R:

$$E_п = -mg\frac{R^2}{r}. \tag{8.6}$$

Если высота тела над поверхностью Земли h мала по сравнению с радиусом Земли ($r=R+h$, $h \ll R$), то формулу (8.6) можно упростить, воспользовавшись тем, что $1/(1+x) \approx \approx 1-x$ (при $x \ll 1$):

$$E_п = -\frac{mgR^2}{R+h} = -\frac{mgR}{1+h/R} \approx -mgR + mgh. \tag{8.7}$$

Первое слагаемое в правой части (8.7) можно опустить, так как оно постоянно, т. е. не зависит от положения тела. Тогда мы получаем

$$E_п = mgh,$$

что, в отличие от (8.6), соответствует выбору начала отсчета потенциальной энергии на поверхности Земли.

Потенциальная энергия кулоновского взаимодействия двух точечных зарядов q_1 и q_2 на расстоянии r друг от друга дается выражением, аналогичным (8.5):

$$E_п = \frac{1}{4\pi\varepsilon_0} \frac{q_1 q_2}{r}. \quad (8.8)$$

Различие знаков в формулах (8.5) и (8.8) вызвано тем, что гравитационное взаимодействие всегда имеет характер притяжения, в то время как при электростатическом взаимодействии заряды одного знака отталкиваются. В случае сил притяжения потенциальная энергия возрастает при увеличении расстояния между телами, а в случае отталкивания — убывает. Обобщение формулы (8.8) на случай системы взаимодействующих зарядов будет проведено в § 4 раздела 3. При упругой деформации, описываемой законом Гука, потенциальная энергия деформированного тела пропорциональна квадрату величины деформации. Например, энергия растянутого стержня, для которого закон Гука записывается в виде (6.9), равна

$$E_п = \frac{1}{2} k (\Delta l)^2. \quad (8.9)$$

Установим соотношение между силой и изменением потенциальной энергии. Рассмотрим перемещение тела $\Delta \boldsymbol{l}$ между двумя близкими точками. Работа сил поля при этом перемещении равна $\boldsymbol{F}\,\Delta \boldsymbol{l}$. С другой стороны, эта работа равна разности потенциальных энергий тела в начальной и конечной точках, т. е. взятому с обратным знаком изменению потенциальной энергии. Поэтому

$$\boldsymbol{F}\,\Delta \boldsymbol{l} = -\Delta E_п.$$

Левую часть этого соотношения можно записать в виде произведения проекции силы F_l на направление перемещения $\Delta \boldsymbol{l}$ на величину этого перемещения Δl: $\boldsymbol{F}\,\Delta \boldsymbol{l} = F_l \Delta l$. Отсюда

$$F_l = -\frac{\Delta E_п}{\Delta l}. \quad (8.10)$$

Проекция силы на произвольное направление может быть найдена делением изменения потенциальной энергии при малом перемещении вдоль этого направления на величину перемещения.

Потенциальная энергия частицы в силовом поле является функцией ее координат. Приравнивая $E_п(x, y, z)$ постоянной величине, получаем уравнение поверхности, во всех точках которой потенциальная энергия имеет одно и то же значение. Эти поверхности постоянной потенциальной энергии дают наглядную картину силового поля и широко используются в электростатике. Сила в каждой точке направлена перпендикулярно к проходящей через эту точку эквипотенциальной поверхности. Это легко увидеть с помощью формулы (8.10). В самом деле, выберем перемещение $\Delta \boldsymbol{l}$ вдоль поверхности постоянной энергии. Тогда $\Delta E_п = 0$ и, следовательно, равна нулю проекция силы на поверхность $E_п = \text{const}$. Так, например, в гравитационном поле, создаваемом телом массы m_1 со сферически симметричным распределением масс, потенциальная энергия тела массы m_2 дается выражением (8.5). Поверхности постоянной энергии такого поля представляют собой сферы, центры которых совпадают с силовым центром. Действующая на массу m_2 сила перпендикулярна эквипотенциальной поверхности и направлена к силовому центру. Проекцию этой силы на радиус, проведенный из силового центра, можно найти из выражения для потенциальной энергии (8.5) с помощью формулы (8.10):

$$F_r = -\frac{\Delta E_п}{\Delta r} = \frac{\gamma m_1 m_2}{\Delta r}\left(\frac{1}{r+\Delta r} - \frac{1}{r}\right) = -\frac{\gamma m_1 m_2}{r(r+\Delta r)},$$

что при $\Delta r \to 0$ дает

$$F_r = -\gamma \frac{m_1 m_2}{r^2}.$$

Полученный результат подтверждает приведенное выше без доказательства выражение для потенциальной энергии (8.5).

Перейдем теперь к обсуждению закона сохранения механической энергии. Рассмотрим систему материальных точек, на которые действуют внешние силы и внутренние консервативные силы взаимодействия. Изменение кинетической энергии всей системы при переходе ее из состояния *1* в состояние *2* равно сумме работ внешних и внутренних сил, действующих на все тела:

$$E_{к2} - E_{к1} = A_{внеш} + A_{внутр}. \qquad (8.11)$$

Работа внутренних консервативных сил равна разности

значений потенциальной энергии системы в начальном и конечном состояниях:

$$A_{внутр} = E_{п1} - E_{п2}.$$

Подставляя это выражение в (8.11) и перегруппировывая слагаемые, получим

$$(E_{к2} + E_{п2}) - (E_{к1} + E_{п1}) = A_{внеш}. \qquad (8.12)$$

Из этого равенства видно, что работа действующих на систему внешних сил определяет изменение функции состояния системы, равной сумме ее кинетической и потенциальной энергий. Эта сумма и есть полная механическая энергия системы.

Если наряду с консервативными в системе действуют и неконсервативные внутренние силы, например силы трения, то изменение механической энергии системы будет равно сумме работ внешних сил и внутренних неконсервативных сил.

Иногда некоторые из действующих на систему внешних сил также являются консервативными, например, когда система находится во внешнем гравитационном (или электростатическом) поле. В этом случае можно ввести понятие потенциальной энергии системы во внешнем поле и включить ее в механическую энергию системы. Теперь изменение механической энергии будет равно сумме работ только неконсервативных внешних и внутренних сил.

В отсутствие неконсервативных сил механическая энергия системы остается неизменной. Возможны лишь взаимные превращения потенциальной энергии в кинетическую и обратно, но полный запас механической энергии системы измениться не может.

Законы сохранения энергии и импульса выполняются в любых инерциальных системах отсчета. Хотя и механическая энергия, и импульс рассматриваемой системы материальных точек имеют разные значения в разных инерциальных системах отсчета, но при движении эти значения остаются неизменными в замкнутых механических системах, а при наличии внешних воздействий в каждой системе отсчета меняются в соответствии с уравнениями (7.4) и (8.12). При этом механическая энергия, в отличие от импульса, может изменяться из-за наличия внутренних неконсервативных сил.

Законы сохранения энергии и импульса тесно связаны с определенными свойствами симметрии пространства и времени. Хотя выше они были получены как следствие законов динамики Ньютона, в действительности они представляют собой более общие принципы, область их применимости шире и не ограничивается ньютоновской динамикой.

Сохранение импульса в замкнутой системе связано с однородностью пространства. Однородность пространства означает, что все явления в замкнутой системе не изменятся, если осуществить параллельный перенос системы из одного места в другое таким образом, чтобы все тела в ней оказались в тех же условиях, в каких они находились в прежнем положении. При таком переносе потенциальная энергия взаимодействия тел, которая, как это следует из однородности пространства, зависит только от их взаимного расположения, остается неизменной. Значит, при переносе всех тел на один и тот же вектор $\boldsymbol{R}$ равна нулю работа всех внутренних сил $\boldsymbol{F}_i$:

$$\sum_i \boldsymbol{F}_i \boldsymbol{R} = 0.$$

Так как $\boldsymbol{R}$ — произвольный вектор, то отсюда следует, что $\sum_i \boldsymbol{F}_i = 0$, т. е. сумма сил в замкнутой системе равна нулю. Это и есть то условие, при выполнении которого второй закон Ньютона приводит к закону сохранения импульса. В этих рассуждениях третий закон Ньютона уже не используется.

Сохранение энергии в замкнутой системе связано с однородностью времени. Однородность времени заключается в том, что все явления в замкнутой системе при одинаковых начальных условиях будут дальше протекать совершенно одинаково, независимо от того, в какой момент времени эти начальные условия созданы.

§ 9. Столкновения частиц

Законы сохранения энергии и импульса часто позволяют гораздо проще получить ответы на некоторые вопросы, связанные с движением тел, чем непосредственное применение законов динамики. Разумеется, информация, получаемая с помощью законов сохранения, не является такой исчерпывающей, как при использовании законов динамики.

но зато и получается она гораздо более легким путем. Особенно ценным здесь является то обстоятельство, что зачастую законы сохранения могут быть использованы даже в тех случаях, когда действующие силы неизвестны. Так обстоит дело, например, в физике элементарных частиц.

Законы сохранения энергии и импульса фактически являются единственным средством теоретического изучения процессов столкновения тел, когда характер действующих при столкновении сил неизвестен. Под столкновениями в физике понимают самые разнообразные процессы взаимодействия между телами при условии, что на бесконечно большом расстоянии друг от друга тела являются свободными. Когда тела проходят мимо друг друга, они взаимодействуют между собой, и результаты такого взаимодействия могут быть самыми разнообразными: тела могут соединиться вместе в одно тело (абсолютно неупругий удар), в результате соударения могут возникнуть новые тела, может случиться и так, что после взаимодействия тела вновь расходятся без изменения своего внутреннего состояния (абсолютно упругий удар). Столкновения макроскопических тел всегда в той или иной степени являются неупругими, однако в области физики атомных явлений и процессов с элементарными частицами понятие об упругом ударе играет важную роль, так как благодаря дискретному характеру энергетического спектра сталкивающихся частиц их внутреннее состояние либо не меняется вообще (упругий удар), либо скачком изменяется на конечную величину.

Рассмотрим абсолютно неупругий удар двух тел, при котором они соединяются вместе и движутся дальше как одно тело. Слипание вместе пластилиновых шаров, застревание пули в деревянном бруске, захват нейтрона атомным ядром — все это примеры абсолютно неупругого удара. Если сталкивающиеся тела образуют замкнутую систему, в которой действуют только внутренние силы, то полный импульс системы остается неизменным. Это позволяет легко определить скорость тела, образовавшегося в результате неупругого соударения двух тел.

Обозначим скорости тел с массами m_1 и m_2 до удара через $\boldsymbol{v}_1$ и $\boldsymbol{v}_2$, а скорость образовавшегося при неупругом ударе тела массы m_1+m_2 через $\boldsymbol{v}$. Тогда, приравнивая полные импульсы до и после удара:

$$m_1\boldsymbol{v}_1 + m_2\boldsymbol{v}_2 = (m_1 + m_2)\,\boldsymbol{v},$$

получим

$$\boldsymbol{v}=\frac{m_1\boldsymbol{v}_1+m_2\boldsymbol{v}_2}{m_1+m_2}. \qquad (9.1)$$

Легко видеть, что определяемая формулой (9.1) скорость $\boldsymbol{v}$ есть просто скорость движения центра масс сталкивающихся тел, которая, разумеется, в замкнутой системе остается неизменной.

При неупругом ударе кинетическая энергия поступательного движения сталкивающихся тел убывает, частично превращаясь во внутреннюю энергию. Кинетическая энергия поступательного движения тел системы до удара

$$E_{к1}=\frac{m_1v_1^2}{2}+\frac{m_2v_2^2}{2}, \qquad (9.2)$$

после удара

$$E_{к2}=\frac{(m_1+m_2)\,v^2}{2}. \qquad (9.3)$$

Подставляя в (9.3) скорость v из (9.1) и составляя разность кинетических энергий до и после удара, найдем

$$E_{к1}-E_{к2}=\frac{1}{2}\,\mu\,(\boldsymbol{v}_1-\boldsymbol{v}_2)^2, \qquad (9.4)$$

где $\mu=\frac{m_1m_2}{m_1+m_2}$ — так называемая приведенная масса двух тел. По существу (9.3) есть кинетическая энергия поступательного движения системы как целого, которая в замкнутой системе не меняется, ибо остается неизменной скорость центра масс. Поэтому (9.4) можно рассматривать как кинетическую энергию относительного движения сталкивающихся частиц, которая при неупругом столкновении превращается в другие виды энергии, например в тепло.

В отличие от неупругого, при абсолютно упругом ударе сохраняется не только импульс, но и механическая энергия, так как внутреннее состояние сталкивающихся частиц после удара остается таким же, каким оно было до удара. Так как частицы до и после столкновения являются свободными, то потенциальная энергия отсутствует, и сохранение механической энергии означает сохранение кинетической энергии сталкивающихся частиц.

При изучении закономерностей упругого столкновения будем для простоты считать, что одна из частиц (с массой m_2) до столкновения покоится: $\mathbf{v}_2=0$. Назовем эту частицу мишенью, а налетающую частицу с массой m_1 и скоростью $\mathbf{v}_1$ — снарядом. Скорости частиц, разлетающихся после столкновения, обозначим через $\mathbf{v}_1'$ и $\mathbf{v}_2'$. Тогда законы сохранения импульса и энергии запишутся в виде

$$m_1\mathbf{v}_1 = m_1\mathbf{v}_1' + m_2\mathbf{v}_2', \tag{9.5}$$

$$\frac{1}{2}m_1v_1^2 = \frac{1}{2}m_1v_1'^2 + \frac{1}{2}m_2v_2'^2. \tag{9.6}$$

Начнем с некоторых простых частных случаев. Прежде всего рассмотрим «лобовое» столкновение частиц, например шаров, при котором скорость $\mathbf{v}_1$ направлена по линии, соединяющей их центры. Тогда и скорости шаров после удара будут направлены по этой же линии. Проектируя равенство (9.5) на это направление, получим скалярное уравнение, которое вместе с (9.6) образует систему уравнений для нахождения проекций v_1' и v_2' скоростей шаров после удара. Решая ее, находим

$$v_1' = \frac{m_1-m_2}{m_1+m_2}v_1, \quad v_2' = \frac{2m_1}{m_1+m_2}v_1. \tag{9.7}$$

Если массы шаров одинаковы ($m_1=m_2$), то первый шар при ударе останавливается, а второй шар после удара движется с такой же скоростью, как и первый шар до удара. Если снаряд легче мишени ($m_1<m_2$), то согласно (9.7) $v_1'<0$, т. е. снаряд отскакивает назад, причем при $m_1\ll m_2$ скорость снаряда просто меняет свое направление на противоположное. Если снаряд тяжелее мишени, то после удара снаряд продолжает двигаться в том же направлении с меньшей скоростью.

Рассматривая изменение кинетической энергии шаров в результате удара, можно убедиться, что в случае равных масс происходит полный обмен энергией, в то время как при большой разнице в массах снаряд при столкновении может передать мишени лишь малую часть своей энергии. В самом деле, пусть, например, снаряд много легче мишени: $m_1\ll m_2$. Тогда, пренебрегая в знаменателе формулы для v_2' (9.7) величиной m_1 по сравнению с m_2, получим $v_2'\approx 2\frac{m_1}{m_2}v_1$, откуда

для кинетической энергии мишени после удара имеем

$$\frac{m_2 v_2'^2}{2} \approx 4\frac{m_1}{m_2}\frac{m_1 v_1^2}{2} \ll \frac{m_1 v_1^2}{2}.$$

Аналогичный результат получится и в случае. если снаряд значительно тяжелее мишени ($m_1 \gg m_2$).

В действительности лобовой удар — это большая редкость. Его относительно легко осуществить разве что при игре в бильярд, а при столкновениях молекул, атомов и элементарных частиц подавляющее число ударов являются нецентральными.

Если частица налетает на неподвижную частицу такой же массы, то при нецентральном упругом ударе частицы разлетаются под прямым углом друг к другу. Действительно, законы сохранения импульса и энергии (9.5) и (9.6) при $m_1 = m_2$ принимают вид

$$\boldsymbol{v}_1 = \boldsymbol{v}_1' + \boldsymbol{v}_2',$$
$$v_1^2 = v_1'^2 + v_2'^2.$$

Первое из этих равенств означает, что векторы скоростей $\boldsymbol{v}_1$, $\boldsymbol{v}_1'$ и $\boldsymbol{v}_2'$ образуют треугольник, а второе — что для этого треугольника справедлива теорема Пифагора, т. е. он прямоугольный: угол между катетами $\boldsymbol{v}_1'$ и $\boldsymbol{v}_2'$ равен $\pi/2$. Однако законов сохранения энергии и импульса недостаточно для определения направления векторов $\boldsymbol{v}_1'$ и $\boldsymbol{v}_2'$ относительно направления движения налетающей частицы. Для того чтобы определить эти направления, нужно знать закон взаимодействия между частицами и их положение в момент столкновения.

В общем случае частиц с разными массами применение законов сохранения к изучению процесса столкновения удобно интерпретировать геометрически. Для этого перейдем из лабораторной инерциальной системы отсчета, в которой частица-мишень до столкновения покоится, в другую инерциальную систему отсчета, в которой центр масс сталкивающихся частиц покоится как до столкновения, так и после. Эта система отсчета движется относительно лабораторной с такой же скоростью, как и центр масс:

$$\boldsymbol{v} = \frac{m_1 \boldsymbol{v}_1 + m_2 \boldsymbol{v}_2}{m_1 + m_2} = \frac{m_1}{m_1 + m_2}\boldsymbol{v}_1,$$

так как мишень до столкновения в лабораторной системе отсчета покоится ($\boldsymbol{v}_2=0$).

В системе центра масс движутся обе частицы — как снаряд, так и мишень. Их импульсы равны по величине и противоположны по направлению, так что полный импульс сталкивающихся частиц в этой системе отсчета равен нулю.

В силу закона сохранения импульса импульсы обеих частиц остаются равными по величине и противоположными по направлению и после столкновения, а в силу закона сохранения энергии остаются неизменными и их величины.

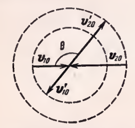


Рис. 9.1. Векторы скоростей частиц до и после столкновения в системе центра масс.

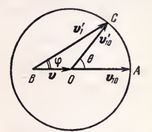


Рис. 9.2. Построение вектора скорости частицы-снаряда после столкновения $\boldsymbol{v}'_1$ в лабораторной системе отсчета.

Тем самым в системе центра масс столкновение сводится к повороту скоростей обеих частиц, остающихся противоположно направленными и неизменными по величине. Это изображено на рис. 9.1, где векторы скоростей с индексом «0» относятся к системе центра масс. Угол θ в зависимости от взаимного расположения частиц при столкновении может принимать любые значения. Его значение не может быть найдено только из законов сохранения.

Скорость частиц в лабораторной системе отсчета можно получить из рис. 9.1 следующим графическим построением. Отложим вектор $\overrightarrow{OA}$, равный скорости снаряда в системе центра масс до удара $\boldsymbol{v}_{10}$ (рис. 9.2). Скорость снаряда в лабораторной системе $\boldsymbol{v}_1$ равна сумме $\boldsymbol{v}_{10}$ и скорости центра масс $\boldsymbol{v}$, т. е. изображается вектором $\overrightarrow{BA}$ на рис. 9.2. После столкновения скорость снаряда в системе центра

масс $\boldsymbol{v}'_{10}$ имеет ту же величину, что и $\boldsymbol{v}_{10}$, и, следовательно, изображается некоторым вектором $\overrightarrow{OC}$, конец которого лежит в какой-то точке окружности радиуса v_{10} с центром в точке O. Поэтому скорость снаряда после удара в лабораторной системе отсчета $\boldsymbol{v}'_1$, равная сумме $\boldsymbol{v}'_{10}$ и скорости центра масс $\boldsymbol{v}$, изображается вектором $\overrightarrow{BC}$. Угол φ характеризует изменение направления скорости снаряда в лабораторной системе в результате столкновения.

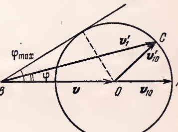


Рис. 9.3. Максимальный угол отклонения частицы-снаряда при столкновении φ_{max}, когда снаряд тяжелее мишени.

Величина вектора $\overrightarrow{BA}$ равна v_1, а вектора $\overrightarrow{BO}$ $v = \frac{m_1}{m_1+m_2} v_1$, поэтому величина вектора $\overrightarrow{OA}$ равна $v_{10} = v_1 - v = \frac{m_2}{m_1+m_2} v_1$. Рис. 9.2 соответствует случаю $m_1 < m_2$, т. е. когда снаряд легче мишени. Так как точка B лежит при этом внутри окружности, то угол отклонения снаряда φ может принимать любое значение. Если снаряд тяжелее мишени ($m_1 > m_2$), то точка B находится вне окружности (рис. 9.3). Видно, что в этом случае угол отклонения снаряда φ не может превышать некоторого максимального значения φ_{max}, синус которого равен отношению v_{10}/v:

$$\sin \varphi_{max} = \frac{v_{10}}{v} = \frac{m_2}{m_1}.$$

Например, при упругом рассеянии дейтронов на неподвижных протонах, когда отношение масс $m_2/m_1 = 1/2$, угол рассеяния не может превышать 30°.

В заключение отметим, что полученные выше закономерности остаются в силе и в тех случаях, когда рассматриваемая система сталкивающихся частиц, строго говоря, не является замкнутой. Необходимо лишь, чтобы внешние силы были малы по сравнению с внутренними силами, действующими во время столкновения. Тогда импульсом внешних сил за время столкновения можно пренебречь.

Например, магнитиое поле в камере Вильсоиа существенно нскрнвляет траекторию заряжеиных частнц до и после столкновеиия, одиако во время столкиовеиия действием магинтного поля можно пренебречь.

§ 10. Законы сохранения и космическне скоростн

Аиалитнческое решенне дниамнческой задачи о двнжеини тела в центральном гравнтацнонном поле, например о двнжении планет вокруг Солица нли нскусственных спутников Землн, сопряжено со зиачнтельнымн математнческимн трудиостямн. Между тем ответы на многне вопросы, касающнеся этого движения, можно сравинтельио просто получить с помощью законов сохраиеиня.

Прежде всего определнм вторую космическую скорость v_{II}, т. е. миннмальиую скорость, которую иужно сообщнть иаходящемуся на поверхностн Земли телу для того, чтобы оно удалнлось на бесконечность (ее часто называют также скоростью освобождення). Проще всего это сделать, нспользуя закон сохранення энергии. Будем счнтать, что двнгателн ракеты срабатывают непосредственно у поверхиостн Землн, сообщают ракете необходнмую скорость и выключаются. Книетнческая энергня тела прн запуске равиа $mv_{II}^2/2$, а потеицнальная энергия вблизи поверхиостн Землн, в соответствнн с формулой (8.6), равиа $-mgR$. Полная механнческая энергня

$$E=\frac{mv_{II}^2}{2}-mgR$$

в свободиом полете остается иеизмеииой. В коиечиом состояинн, когда ракета удалнлась от Землн на бескоиечиость, ее потенцнальная энергня равна нулю. Очевндио, что необходнмая иачальная скорость будет иаименьшей, если в конечном состояннн скорость ракеты обратнтся в иуль. Следовательио, в конечном состоянни полная механическая энергия равиа иулю, и вследствие закона сохраиеиня эиергнн

$$\frac{mv_{II}^2}{2}-mgR=0,$$

откуда иемедлеиио получаем

$$v_{II}=\sqrt{2gR}=11{,}2\ \text{км/с}. \qquad (10.1)$$

Ракета удалится на бесконечность независимо от того, в каком направлении сообщена ей вторая космическая скорость, хотя траектории движения при этом будут, разумеется, разные.

Обратим внимание на то, что приведенные выше рассуждения проведены в системе отсчета, связанной с Землей. Это совершенно естественно, так как по самой постановке задачи очевидно, что Земля считается неподвижной. Точнее, система отсчета, связанная с Землей, считается инерциальной, а притяжением Солнца и тем более других планет пренебрегается. Однако попробуем в том же самом приближении (т. е. пренебрегая притяжением к Солнцу) рассмотреть этот вопрос с точки зрения системы отсчета, связанной с Солнцем. Это совершенно законно, так как гелиоцентрическая система отсчета является инерциальной с большей степенью точности, чем геоцентрическая. Поскольку притяжение к Солнцу здесь не учитывается, Солнце является лишь телом, с которым связана система отсчета, а не физическим телом, влияющим на движение.

Обозначим скорость движения Земли через $\boldsymbol{v}$ и предположим для простоты, что начальная скорость тела относительно Земли $\boldsymbol{v}_{II}$ совпадает с ней по направлению. Применим закон сохранения энергии, учитывая, что, удалившись от Земли на бесконечность, тело останавливается относительно Земли, т. е. в рассматриваемой системе отсчета имеет ту же скорость, что и Земля:

$$\frac{m(v_{II}+v)^2}{2}-mgR=\frac{mv^2}{2}. \qquad (10.2)$$

Отсюда находим

$$v_{II}=\sqrt{2gR+v^2}-v, \qquad (10.3)$$

что не совпадает с полученным ранее значением. Какому же результату верить? Совершенно очевидно, что формула (10.3) не может быть верной. В нее входит v — относительная скорость двух использованных нами систем отсчета. Но так как все инерциальные системы отсчета равноправны, то ответ не может зависеть от v. В чем же дело? В справедливости закона сохранения энергии сомневаться не приходится. Выражение для потенциальной энергии во всех инерциальных системах отсчета одинаково. Значит, в уравнении баланса энергии (10.2) что-то не учтено. Что же? Единст-

вениое, что мы могли упустить,— это изменение кинетической энергии Земли. В самом деле,. при удалении тела от Землн сила тяготения действует не только на тело, ио и на Землю, оказывая влияние иа ее движение. Правда, изменение кинетической энергии Земли при этом очень мало, ибо ее масса M много больше, чем масса тела m. Тем не менее попробуем учесть его аккуратно. Обозначая скорость Земли после удаления тела на бесконечность через v_1, запишем закон сохранения эиергии в виде

$$\frac{Mv^2}{2}+\frac{m(v_{II}+v)^2}{2}-mgR=\frac{Mv_1^2}{2}+\frac{mv_1^2}{2}. \qquad (10.4)$$

Скорость тела в коиечном состоянии теперь равиа v_1, ибо тело, как и раньше, должио быть неподвижно относительио Земли. Величину v_1 можно иайти с помощью закона сохранения импульса, ибо в отсутствие внешних сил, т. е. в замкнутой системе взаимодействующих тел, полный импульс сохраняется. Поэтому

$$Mv+m(v_{II}+v)=Mv_1+mv_1. \qquad (10.5)$$

Находя v_1 из уравнения (10.5) и подставляя в (10.4), после несложных преобразований получим

$$v_{II}^2=\left(1+\frac{m}{M}\right)2gR. \qquad (10.6)$$

Это выражеиие уже гораздо ближе к прежиему результату, чем (10.3), ио все-таки отличается от него лишним множителем $1+m/M$. Заметим, что если считать массу тела много меньше массы Земли, т. е. $m/M\ll 1$, то слагаемым m/M можно пренебречь по сравнению с единнцей, и формула (10.6) дает прежний результат $v_{II}=\sqrt{2gR}$. Нетрудно сообразить, что предположение $m/M\to 0$ с самого начала было иеявно использовано при решении задачн в системе отсчета, связанной с Землей. Действительно, Земля считалась неподвижной как в иачальиом, так н в конечном состоянии, несмотря на то, что на иее, как и на тело, действовала сила тяготения. Это возможно, только если $m/M\to 0$.

Итак, выражение (10.6) является более общим, чем (10.1). Оно дает возможность определить вторую космическую скорость в том случае, когда массы запускаемого тела и Землн сравиимы между собой. Однако теперь возникает другой вопрос. Почему преиебрежеиие измеиеиием кинетической

энергии Земли (при $m/M \to 0$) в геоцентрической системе отсчета допустимо, а в гелиоцентрической приводит к явно неверному результату (10.3)? Ведь изменение скорости Земли одинаково в любой инерциальной системе отсчета. В этом легко убедиться, переписав формулу (10.5) в несколько ином виде:

$$mv_{II} = (M + m)(v_1 - v). \qquad (10.7)$$

Видно, что изменение скорости Земли $\Delta v = v_1 - v$ не зависит от ее начальной скорости v, т. е. от выбора системы отсчета. Однако изменение кинетической энергии Земли в разных системах отсчета будет разным: в геоцентрической системе это $M(\Delta v)^2/2$, а в гелиоцентрической

$$\frac{M(v+\Delta v)^2}{2} - \frac{Mv^2}{2} = Mv\,\Delta v + \frac{M(\Delta v)^2}{2}. \qquad (10.8)$$

При $m \ll M$, как видно из (10.7), $\Delta v \ll v_{II}$. Поскольку скорость v больше второй космической скорости, то первое слагаемое в правой части (10.8) много больше второго, т. е. изменение кинетической энергии в гелиоцентрической системе много больше этого изменения в геоцентрической, и его нельзя не учитывать в уравнении баланса энергии.

Разумеется, формула (10.6) может быть получена и в геоцентрической системе отсчета, если там учесть изменение кинетической энергии Земли.

Разобранный пример очень поучителен. Он наглядно показывает, с какой осторожностью нужно подходить к вопросу о том, что́ существенно в рассматриваемом явлении, а чем можно пренебречь. Все инерциальные системы отсчета равноправны в том смысле, что законы природы в них одинаковы. Поэтому при точном решении задачи выбор такой системы безразличен. Однако при нахождении приближенного решения пренебрежения, допустимые в одной системе отсчета, могут оказаться совершенно непригодными в другой.

Перейдем теперь к определению третьей космической скорости v_{III}, т. е. минимальной скорости, которую нужно сообщить телу вблизи поверхности Земли для того, чтобы оно смогло покинуть пределы Солнечной системы. Будем рассуждать следующим образом. Забудем на время о земном тяготении и найдем минимальную скорость v_1, которую нужно сообщить телу, находящемуся от Солнца на расстоя-

нии r, равном радиусу земной орбиты, чтобы оно смогло преодолеть притяжение Солнца. Эту скорость легко найти, используя закон сохранения энергии. Поскольку мы пока пренебрегаем полем тяготения Земли, то нужно просто потребовать, чтобы сумма кинетической энергии тела $mv_1^2/2$ и потенциальной энергии в поле тяготения Солнца $-\gamma m M_C/r$ равнялась нулю: тело должно остановиться на бесконечно большом расстоянии от Солнца, где потенциальная энергия обращается в нуль. Отсюда

$$v_1^2 = 2\gamma \frac{M_C}{r}. \qquad (10.9)$$

Легко видеть, что эта скорость в $\sqrt{2}$ раз больше скорости Земли на круговой орбите движения вокруг Солнца $v = 29,8$ км/с. Тело удалится на бесконечность независимо от того, в каком направлении сообщена ему скорость v_1. Итак, v_1 равна приблизительно 42,1 км/с. Это очень много, однако, разумеется, мы можем использовать движение Земли и запускать тело в ту же сторону, куда движется Земля по орбите. Тогда телу нужно сообщить добавочную скорость, равную $(\sqrt{2}-1)v \approx 12,3$ км/с.

Теперь нетрудно найти и саму третью космическую скорость. Для этого достаточно только сообразить, что на самом деле скорость 12,3 км/с тело должно иметь после того, как оно преодолеет притяжение к Земле. Поэтому сумма кинетической энергии тела при запуске $mv_{III}^2/2$ и потенциальной энергии на поверхности Земли $-mgR$ должна равняться кинетической энергии движения со скоростью 12,3 км/с после преодоления земного тяготения:

$$\frac{mv_{III}^2}{2} - mgR = \frac{m(\sqrt{2}-1)^2 v^2}{2},$$

откуда

$$v_{III}^2 = (\sqrt{2}-1)^2 v^2 + 2gR. \qquad (10.10)$$

Этой формуле можно придать другой вид, если вспомнить, что $\sqrt{2gR}$ равен второй космической скорости v_{II}:

$$v_{III}^2 = (\sqrt{2}-1)^2 v^2 + v_{II}^2. \qquad (10.11)$$

Подставляя сюда численные значения орбитальной скорости Земли $v \approx 29,8$ км/с и второй космической скорости $v_{II} \approx 11,2$ км/с, получим $v_{III} \approx 16,7$ км/с.

Итак, ответ получен. Но у читателя, возможно, возник вопрос: почему рассуждения проводились в два этапа? Другими словами, почему закон сохранения энергии использовался дважды — сначала для процесса выхода тела из поля тяготения Солнца, а затем — из поля тяготения Земли? Нельзя ли применить закон сохранения энергии один раз ко всему процессу в целом, потребовав, чтобы полная энергия тела, т. е. сумма его кинетической энергии и потенциальных энергий в полях тяготения Земли и Солнца, равнялась нулю:

$$\frac{m(v+v_{III})^2}{2}-mgR-\gamma\frac{mM_C}{r}=0? \qquad (10.12)$$

Однако читатель, разобравшийся в предыдущем примере, конечно, сообразил, что так писать нельзя. Действительно, выразив второе слагаемое в формуле (10.12) через вторую космическую скорость $v_{II}^2=2gR$, а третье — через скорость Земли на круговой орбите вокруг Солнца $v^2=\gamma M_C/r$, мы не получим для третьей космической скорости формулы (10.11). И совершенно понятно — почему: в выражении (10.12) мы не учитывали изменения кинетической энергии Земли при удалении от нее запущенного тела. Хотя это изменение и мало, но, как мы видели в предыдущем примере, учет его в гелиоцентрической системе отсчета необходим. Учтем изменение кинетической энергии Земли. Разумеется, при этом мы будем пренебрегать изменением кинетической энергии Солнца: как при вычислении второй космической скорости можно было пренебречь изменением кинетической энергии Земли при использовании связанной с ней системы отсчета, так и здесь изменением кинетической энергии Солнца можно пренебречь при использовании гелиоцентрической системы отсчета. Его нужно было бы учитывать, если бы мы использовали какую-нибудь инерциальную систему отсчета, в которой Солнце движется, например систему отсчета, связанную с какой-либо галактикой. С учетом сказанного закон сохранения энергии в гелиоцентрической системе отсчета следует писать в виде

$$\frac{m(v+v_{III})^2}{2}+\frac{Mv^2}{2}-mgR-\gamma\frac{mM_C}{r}=\frac{Mv_2^2}{2}. \qquad (10.13)$$

В этом выражении M — масса Земли, v_2 — скорость Земли после удаления тела, а остальные обозначения прежние.

Третье и четвертое слагаемые в левой части, как и раньше, выразим соответственно через вторую космическую скорость и скорость Земли на круговой орбите. Перенесем второе слагаемое из левой части уравнения (10.13) в правую; тогда в правой части будет стоять изменение кинетической энергии Земли, которое представим в виде

$$\frac{M}{2}(v_2^2-v^2)=\frac{M}{2}(v_2+v)(v_2-v)\approx Mv(v_2-v). \quad (10.14)$$

Поскольку масса тела много меньше массы Земли, то изменение скорости Земли при удалении тела от нее мало, и сумма v_2+v приближенно заменена на $2v$. Нетрудно сообразить, что это соответствует пренебрежению вторым слагаемым в правой части формулы (10.8).

Для нахождения изменения скорости Земли $\Delta v=v_2-v$ воспользуемся законом сохранения импульса. Пренебрежем влиянием поля тяготения Солнца на движение Земли и запущенного тела в течение всего времени, которое оно затрачивает на выход из зоны действия земного тяготения. Так как скорость тела при выходе из этой зоны равна v_1, имеем

$$m(v+v_{III})+Mv=mv_1+Mv_2.$$

Отсюда

$$M(v_2-v)=m(v+v_{III}-v_1)=m\left[v_{III}-(\sqrt{2}-1)\,v\right], \quad (10.15)$$

так как $v_1=\sqrt{2}\,v$. Обратим внимание на то, что изменение скорости Земли Δv стремится к нулю при $m/M\to 0$, т. е. запуск космических аппаратов практически не влияет на движение самой Земли. Умножая (10.15) на v, получим, согласно (10.14), изменение кинетической энергии Земли. Подставляя это изменение в уравнение баланса энергии (10.13), получим уравнение для определения третьей космической скорости v_{III}:

$$(v+v_{III})^2-2v_{II}^2-2v^2=2v\left[v_{III}-(\sqrt{2}-1)\,v\right].$$

Решая это уравнение, находим для v_{III} прежнее значение, даваемое формулой (10.11).

Обратим внимание на следующее обстоятельство. Несмотря на то, что закон сохранения энергии (10.13) был записан для всего процесса в целом, при нахождении изменения скорости Земли нам пришлось воспользоваться зако-

ном сохранения импульса в приближенном виде только для определенного этапа процесса, а именно для выхода тела только из зоны действия тяготения Земли. При этом мы считали, что на втором этапе, т. е. при удалении тела из зоны действия солнечного притяжения, скорость Земли уже не менялась по величине. Таким образом, фактически нам все равно пришлось проводить поэтапное приближенное рассмотрение. Попытка применить закон сохранения импульса ко всему процессу не привела бы к желаемому результату. Дело в том, что здесь мы сталкиваемся с так называемой «задачей трех тел», движущихся под действием сил взаимного притяжения. Точное решение этой задачи в общем случае встречает колоссальные математические трудности и может быть доведено до конца лишь в некоторых частных случаях.

При решении практических задач космической динамики обычно используется приближенный подход, основанный на разбиении пространства на так называемые сферы действия отдельных небесных тел. Так, например, в разобранном примере сначала рассматривалось движение тела только под действием притяжения к Земле. При этом, строго говоря, пренебрегается не влиянием Солнца на движение тела, а разностью во влияниях Солнца на движения Земли и тела, т. е. фактически пренебрегается неоднородностью поля тяготения Солнца в сфере действия Земли. После выхода тела из сферы действия Земли рассматривалось его движение только в поле тяготения Солнца. Размер сферы действия Земли определяется тем расстоянием, на котором разность ускорений, сообщаемых Солнцем Земле и запущенному телу, становится сравнимой с ускорением, сообщаемым телу Землей. В отличие от сферы действия, «сфера притяжения Земли относительно Солнца», определяемая как область, на границе которой равны по величине гравитационные ускорения тела от Земли и от Солнца, не играет никакой роли в космической динамике.

§ 11. Простые примеры из космической динамики

Особенности движения тел в ньютоновском поле тяготения лучше всего изучать, рассматривая конкретные примеры. В этом параграфе будет рассмотрено несколько примеров из космической динамики движения спутников в

гравитационном поле Земли на основе законов Кеплера и законов сохранения.

Пример 1. С полюса Земли запускают две ракеты, одну вертикально вверх, другую горизонтально. Начальные скорости обеих ракет равны v_0, причем величина v_0 больше первой космической скорости и меньше второй. Выясним, какая из ракет удалится дальше от центра Земли и во сколько раз. Сопротивлением воздуха будем пренебрегать.

Рассмотрим вначале более простой случай, когда ракета запускается вертикально вверх. Поскольку единственная сила, действующая на ракету в свободном полете, есть сила притяжения к Земле, направленная вертикально вниз, то ракета полетит по прямой, проходящей через центр Земли. Так как начальная скорость ракеты меньше второй космической скорости, то ракета на некотором расстоянии r_1 от центра Земли остановится и начнет падать назад. Точку максимального удаления проще всего найти из энергетических соображений. Действительно, так как полная механическая энергия системы ракета — Земля сохраняется, энергия в начале полета $(mv_0^2/2-mgR)$ равна энергии в точке остановки $(-mgR^2/r_1)$. Отсюда сразу находим расстояние максимального удаления от центра Земли:

$$r_1=\frac{2gR^2}{2gR-v_0^2}.$$

Прежде чем вычислять величину максимального удаления ракеты при горизонтальном запуске, выясним вопрос о форме траектории. Поскольку начальная скорость ракеты превышает первую космическую, но меньше второй, ракета движется по эллипсу, у которого фокус находится в центре Земли, а начальная точка полета является перигеем. Большая ось эллипса проходит через эту точку и центр Земли (рис. 11.1). Интересующая нас точка наибольшего удаления от центра Земли — апогей — лежит на противоположном конце большой оси, и скорость ракеты v_2 в этой точке, разумеется, отлична от нуля и направлена перпендикулярно большой оси эллипса.

Для нахождения величины r_2 опять можно воспользоваться законом сохранения энергии:

$$\frac{mv_0^2}{2}-mgR=\frac{mv_2^2}{2}-\frac{mgR^2}{r_2}.$$

Уже отсюда легко увидеть, что максимальное удаление ракеты r_2 в этом случае будет меньше, чем r_1. Уравнение содержит две неизвестные величины v_2 и r_2 и поэтому имеет бесчисленное множество решений. Что бы это могло означать? Перечитав еще раз наши рассуждения, легко заметить, что в уравнение закона сохранения энергии не вошли никакие признаки, которые характеризовали бы точку r_2 как точку наибольшего удаления. Точно такое же уравнение мы получили бы и для любой другой точки траектории. Заметим, что в первом случае (при вертикальном запуске ракеты) точка максимального удаления была уже выделена в уравнении закона сохранения энергии, так как только в этой точке кинетическая энергия ракеты обращается в нуль. Какое же условие следует добавить к уравнению баланса энергии во втором случае, чтобы учесть особенности точки наибольшего удаления, отличающие ее от всех других точек траектории? Мы уже заметили, что в этой точке скорость перпендикулярна к направлению на центр Земли. Точно таким же свойством обладает и начальная точка траектории: по условию начальная скорость ракеты v_0 перпендикулярна направлению на центр Земли. Во всех остальных точках траектории это не так. Этот факт позволяет в простом виде применить второй закон Кеплера о постоянстве секторной скорости при движении в центральном поле:

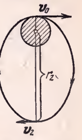


Рис. 11.1. Эллиптическая траектория ракеты при $v_I < v_0 < v_{II}$.

$$v_0 R = v_2 r_2.$$

Теперь мы имеем систему уравнений для определения v_2 и r_2, причем из наших рассуждений вытекает, что эта система должна иметь два решения, соответствующих перигею и апогею. Легко убедиться, что после подстановки $v_2 = v_0 R / r_2$ уравнение баланса энергии превращается в квадратное уравнение относительно r_2, корни которого равны R и $r_2 = v_0^2 R / (2gR - v_0^2)$; сравнивая r_2 с r_1, получаем

$$\frac{r_1}{r_2} = \frac{2gR}{v_0^2} = \left(\frac{v_{II}}{v_0}\right)^2 > 1.$$

Пример 2. Два искусственных спутника Земли движутся по одной и той же круговой орбите на расстоянии L друг от друга, считая вдоль траектории. Период обращения спутников равен T. Как можно сблизить спутники на орбите, если на одном из них есть двигатель, с помощью которого можно практически мгновенно сообщить некоторую дополнительную скорость Δv, малую по сравнению с орбитальной скоростью и направленную по касательной к траектории? Как скорость Δv связана с L и T?

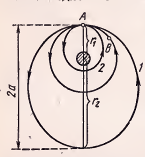


Рис. 11.2. Преобразование круговой траектории в эллиптические при сообщении спутнику в точке A дополнительной скорости в направлении орбитального движения (1) и против движения (2).

На рис. 11.2 показана круговая траектория радиуса r_1, по которой движутся спутники на расстоянии L друг от друга. Если в произвольной точке A спутнику сообщить дополнительную скорость Δv в направлении орбитального движения, то он будет двигаться по эллиптической орбите *1* с фокусом в центре Земли. Период обращения по такой орбите больше, чем по круговой. Если же дополнительную скорость Δv в точке A направить против движения по орбите, то спутник перейдет на эллиптическую орбиту *2*, период обращения по которой меньше T.

Отсюда становится ясным, какие маневры следует проводить для сближения спутников. Прежде всего отметим, что, поскольку круговая и эллиптическая орбиты имеют только одну общую точку A, встреча спутников может произойти только в этой точке. Поэтому встреча может произойти только через промежуток времени после совершения маневра, кратный периоду обращения спутника по эллиптической орбите. Если двигатель установлен на спутнике, идущем впереди, то его нужно разогнать, переводя на орбиту *1*. Если же двигатель находится на спутнике, идущем сзади, то его нужно тормозить, переводя на орбиту *2*. Обратите внимание на кажущийся парадокс: для того чтобы догнать, нужно притормозить.

Для определенности рассмотрим первый случай: в момент совершения маневра пассивный спутник находится в точке B,

отстоящей на L от точки A. Найдем величину дополнительной скорости Δv, которую нужно сообщить активному спутнику, чтобы встреча произошла через один оборот. Обозначим через v скорость движения по круговой орбите. Тогда период обращения T_1 по эллиптической орбите должен быть больше периода обращения T по круговой на время прохождения дуги L по круговой орбите:

$$T_1 - T = \frac{L}{v}. \tag{11.1}$$

Найдем связь между периодом обращения T_1 по эллиптической орбите и той добавочной скоростью Δv, которая переводит спутник на эту орбиту. В этом нам помогут законы Кеплера. На основании третьего закона Кеплера

$$T_1^2 = T^2 \frac{a^3}{r_1^3}, \tag{11.2}$$

где $a = (r_1 + r_2)/2$ — большая полуось эллипса (рис. 11.2).

Для того чтобы связать расстояние r_2 от центра Земли до апогея эллиптической орбиты с Δv, воспользуемся вторым законом Кеплера и законом сохранения энергии: обозначив через $v_1 = v + \Delta v$ скорость спутника в перигее, а через v_2 — в апогее эллиптической орбиты, имеем

$$r_1 v_1 = r_2 v_2, \tag{11.3}$$

$$\frac{mv_1^2}{2} - \frac{mgR^2}{r_1} = \frac{mv_2^2}{2} - \frac{mgR^2}{r_2}. \tag{11.4}$$

Здесь m — масса спутника, R — радиус Земли. Выражая v_2 из (11.3) и подставляя в (11.4), получаем

$$\frac{1}{2} v_1^2 \left(1 - \frac{r_1^2}{r_2^2}\right) = \frac{gR^2}{r_1}\left(1 - \frac{r_1}{r_2}\right). \tag{11.5}$$

Замечая, что gR^2/r_1 есть квадрат скорости спутника на круговой орбите радиуса r_1, из уравнения (11.5) получаем

$$r_2 = \frac{r_1}{2\dfrac{v^2}{v_1^2} - 1}.$$

Отсюда большая полуось эллипса

$$a = \frac{1}{2}(r_1 + r_2) = \frac{r_1}{2 - \dfrac{v_1^2}{v^2}}.$$

Поскольку по условию добавочная скорость Δv много меньше скорости орбитального движения v, то

$$\frac{v_1^2}{v^2}=\left(1+\frac{\Delta v}{v}\right)^2\approx 1+2\frac{\Delta v}{v}.$$

Теперь для большой полуоси имеем $a\approx r_1(1+2\Delta v/v)$. Подставим полученное выражение для a в уравнение третьего закона Кеплера (11.2). Возводя отношение a/r_1 в куб и учитывая, что $\Delta v/v\ll 1$, находим $T_1^2\approx T^2(1+6\Delta v/v)$, откуда

$$T_1-T=\frac{3T\,\Delta v}{v}.$$

Сравнивая это выражение с (11.1), видим, что

$$\Delta v=\frac{L}{3T}. \qquad (11.6)$$

Итак, если мы хотим, чтобы встреча спутников произошла через один оборот, то идущему впереди спутнику нужно сообщить добавочную скорость Δv, определяемую соотношением (11.6).

Чтобы встреча произошла через n оборотов, нужно сообщить в n раз меньшую добавочную скорость $\Delta v=L/(3nT)$. Таким образом, мы можем израсходовать тем меньше топлива для совершения маневра, чем больше времени согласны ждать встречи.

После завершения маневра сближения, для того чтобы перевести спутник снова на круговую орбиту, придется погасить сообщенную ему добавочную скорость, т. е. еще раз включить двигатель.

Второй случай, когда требуется «догнать» идущий впереди спутник, рассматривается совершенно аналогично и приводит к такому же выражению (11.6) для Δv.

Пример 3. На большом расстоянии от Земли метеорит движется относительно нее со скоростью v_0. Если бы земное притяжение отсутствовало, метеорит прошел бы на расстоянии l от центра Земли (рис. 11.3, *а*). Выясним, при каком наибольшем значении «прицельного» расстояния l метеорит будет захвачен Землей.

На большом расстоянии от Земли, где потенциальную энергию взаимодействия с Землей можно считать равной нулю, метеорит имеет скорость v_0 и его полная энергия равна кинетической $mv_0^2/2$. Если бы начальная скорость метео-

рита v_0 была равна нулю, то, двигаясь только под действием силы притяжения к Земле, он обязательно упал бы на Землю и при падении имел у поверхности Земли скорость, равную второй космической $v_{II}=11,2$ км/с, в чем легко убедиться с помощью закона сохранения энергии. Ясно, что

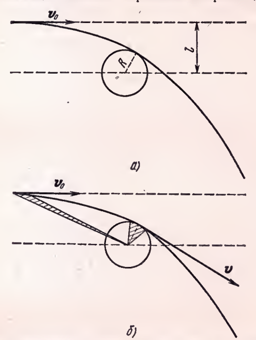


Рис. 11.3. *а)* Траектория метеорита вблизи Земли и «прицельное» расстояние *l*. *б)* Заштрихованные площади равны по второму закону Кеплера.

траектория метеорита в этом случае — прямая, проходящая через центр Земли. Если же начальная скорость метеорита отлична от нуля, то он в поле земного тяготения движется по гиперболе и будет захвачен Землей только тогда, когда эта гипербола «заденет» земной шар. Нетрудно сообразить, что при заданном прицельном расстоянии l траектория метеорита будет тем меньше искривлена, чем больше его скорость v_0, так что достаточно быстрые метеориты благополучно минуют Землю. Очевидно, что наименьшей

скорости v_0, при которой метеорит еще «проскочит» Землю, соответствует траектория, изображенная на рис. 11.3, *а*. И наоборот, при заданной начальной скорости v_0 эта траектория соответствует наибольшему прицельному расстоянию l, при котором метеорит будет захвачен Землей. Итак, для получения ответа на поставленный вопрос нужно рассмотреть траекторию, касающуюся земного шара.

При движении метеорита в поле тяжести Земли выполняется закон сохранения механической энергии

$$\frac{mv_0^2}{2}=\frac{mv^2}{2}-mgR, \qquad (11.7)$$

где v — скорость метеорита в точке касания, отстоящей от центра Земли на расстояние R. Второй закон Кеплера о постоянстве секторной скорости при движении тела в поле тяжести справедлив и для разомкнутых траекторий. Поэтому приравняем секторные скорости метеорита в бесконечно удаленной от Земли точке и в точке касания:

$$lv_0=Rv. \qquad (11.8)$$

Правая часть этого равенства очевидна, поскольку в точке касания вектор скорости $\boldsymbol{v}$ перпендикулярен радиусу Земли, а левая часть становится очевидной, если посмотреть на рис. 11.3, *б*. Подставляя v из уравнения (11.7) закона сохранения энергии в (11.8), находим l:

$$l=R\sqrt{1+\frac{2gR}{v_0^2}}=R\sqrt{1+\left(\frac{v_{11}}{v_0}\right)^2}. \qquad (11.9)$$

Из полученного ответа видно, что максимальное прицельное расстояние, при котором метеорит будет захвачен Землей, зависит от величины начальной скорости v_0. Если $v_0\to 0$, то $l\to\infty$, т. е. первоначально покоившийся относительно Земли метеорит упадет на Землю при любых обстоятельствах (разумеется, это справедливо в предположении, что Солнце и другие планеты практически не влияют на движение метеорита). Если $v_0\to\infty$ ($v_0\gg v_{11}$), то $l\to R$, т. е. в пределе бесконечно большой скорости v_0 траектория прямолинейна, так как за малое время полета метеорита вблизи Земли сила земного притяжения не успевает вызвать заметного изменения импульса метеорита (т. е. искривить его траекторию), и метеорит попадет на Землю только тогда,

когда его прицельное расстояние l не превосходит радиуса Земли.

В приведенном примере не учитывалось влияние земной атмосферы на траекторию метеорита. Однако при расчете максимального прицельного расстояния по формуле (11.9) мы не получим заметной ошибки, так как толщина атмосферы мала по сравнению с радиусом Земли R.

§ 12. Механическое равновесие

Раздел механики, в котором изучаются условия равновесия тел, называется статикой. Проще всего рассмотреть условия равновесия абсолютно твердого тела, т. е. такого тела, размеры и форму которого можно считать неизменными. Понятие абсолютно твердого тела является абстракцией, поскольку все реальные тела под влиянием приложенных к ним сил в той или иной степени деформируются, т. е. меняют свою форму и размеры. Величина деформаций зависит как от приложенных к телу сил, так и от свойств самого тела — его формы и свойств материала, из которого оно изготовлено. Во многих практически важных случаях деформации бывают малыми и использование представлений об абсолютно твердом теле является оправданным.

Однако не всегда малость деформаций является достаточным условием для того, чтобы тело можно было считать абсолютно твердым. Чтобы пояснить это, рассмотрим следующий пример. Балка, лежащая на двух опорах (рис. 12.1, *а*), может рассматриваться как абсолютно твердое тело, несмотря на то, что она слегка прогибается под действием сил тяжести. Действительно, в этом случае условия механического равновесия позволяют определить силы реакции опор $\boldsymbol{N}_1$ и $\boldsymbol{N}_2$, не учитывая деформации балки. Но если та же балка лежит на трех опорах (рис. 12.1, *б*), то представление об абсолютно твердом теле является неприменимым. В самом деле, пусть крайние опоры находятся на одной горизонтали, а средняя — чуть ниже. Если балка абсолютно твердая, т. е. вообще не прогибается, то она совсем не давит на среднюю опору ($N_3=0$). Если же балка прогибается, то она давит на среднюю опору, причем тем сильнее, чем больше деформация. Условия равновесия абсолютно твердого тела в этом случае не позволяют определить силы реакции опор $\boldsymbol{N}_1$, $\boldsymbol{N}_2$ и $\boldsymbol{N}_3$, так как приводят к двум уравнениям для трех неизвестных величин. Такие системы носят название

статическн неопределнмых. Для нх расчета необходимо учнтывать упругне свойства тел.

Прнведенный пример показывает, что прнменнмость моделн абсолютно твердого тела в статнке определяется не

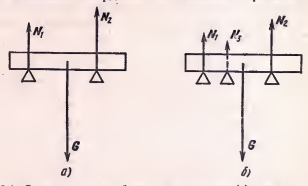


Рис. 12.1. Силы реакцни для балки на двух опорах (а) и на трех опорах (б). В случае б модель абсолютно твердого тела к балке неприменима.

столько свойствамн самого тела, сколько условнямн, в которых оно находнтся.

Условня равновесня абсолютно твердого тела представляют собой частный случай днпамнческнх уравнений, когда ускоренне отсутствует, хотя нсторнческн статнка возннкла нз потребностей стронтельной техннкн почтн на два тысячелетня раньше днпамнкн. В ннерцнальной снстеме отсчета твердое тело находнтся в равновеснн, еслн векторная сумма всех действующих на тело внешннх снл н векторная сумма моментов этнх снл равны нулю. Прн выполненнн первого условия равно нулю ускоренне центра масс тела. Прн выполненнн второго условня отсутствует угловое ускоренне вращення. Поэтому, еслн в начальный момент тело покоилось, то оно будет оставаться в покое н дальше.

В дальнейшем мы огранНчнмся нзученнем сравнительно простых снстем, в которых все действующне снлы лежат в одной плоскости. В этом случае векторное условне

$$\sum_i \boldsymbol{F}_i = 0 \qquad (12.1)$$

сводится к двум скалярным:

$$\sum_i F_{ix} = 0, \quad \sum_i F_{iy} = 0, \qquad (12.2)$$

если расположить оси x и y в плоскости действия сил. Некоторые из входящих в условия равновесия (12.1) действующих на тело внешних сил могут быть заданы, т. е. их величина и направление известны. Что же касается сил реакции связей или опор, ограничивающих возможное перемещение тела, то они, как правило, наперед не заданы и сами подлежат определению. В отсутствие трения силы реакции перпендикулярны поверхности соприкосновения тел.

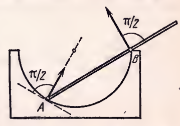


Рис. 12.2. Направление сил реакции, действующих на стержень.

Иногда возникает сомнение в определении направления силы реакции связи, как, например, на рис. 12.2, где изображен стержень, опирающийся концом A о гладкую вогнутую поверхность чашки и в точке B на острый край чашки. Для определения направления сил реакции в этом случае можно мысленно немного подвинуть стержень, не нарушая его контакта с чашкой. Сила реакции будет направлена перпендикулярно поверхности, по которой скользит точка контакта. Так, в точке A действующая на стержень сила реакции перпендикулярна поверхности чашки, а в точке B — перпендикулярна стержню.

Для плоской системы сил векторы моментов всех сил направлены перпендикулярно плоскости, в которой лежат силы, если моменты рассматриваются относительно точки, лежащей в этой же плоскости. Поэтому векторное условие для моментов сводится к одному скалярному: в положении равновесия алгебраическая сумма моментов всех внешних действующих сил равна нулю. Величина момента силы $\boldsymbol{F}$ относительно точки O равна произведению величины силы F на расстояние от точки O до линии, вдоль которой действует сила $\boldsymbol{F}$. При этом моменты, стремящиеся повернуть тело по часовой стрелке, берутся с одним знаком, против часовой стрелки — с противоположным. Выбор точки, относительно которой рассматриваются моменты сил, производится исключительно из соображений удобства: уравнение моментов будет тем проще, чем больше сил будут иметь равные нулю моменты.

Для иллюстрации применения условий равновесия абсолютно твердого тела рассмотрим следующий пример. Легкая лестница-стремянка состоит из двух одинаковых частей, шарнирно соединенных вверху и связанных веревкой у основания (рис. 12.3, *а*). Определим, каково натяжение веревки, с какими силами взаимодействуют половинки лестницы в шарнире и с какими силами они давят на пол, если на середине одной из них стоит человек весом G. Рассматриваемая система состоит из двух твердых тел — половинок лестницы, и условия равновесия можно применять как для системы в целом, так и для ее частей.

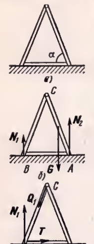


Рис. 12.3. Шарнирно соединенные вверху половинки лестницы-стремянки связаны горизонтальной веревкой у основания.

Применяя условия равновесия ко всей системе в целом, можно найти силы реакции пола N_1 и N_2 (рис. 12.3, *б*). При отсутствии трения эти силы направлены вертикально вверх и условие равенства нулю векторной суммы внешних сил (12.1) принимает вид

$$N_1 + N_2 = G. \qquad (12.3)$$

Условие равновесия моментов внешних сил относительно точки A записывается следующим образом:

$$N_1 2l \cos\alpha = G \frac{l}{2} \cos\alpha, \qquad (12.4)$$

где l — длина половинки лестницы, а α — угол, образованный лестницей с полом. Решая систему уравнений (12.3) и (12.4), находим $N_1 = G/4$, $N_2 = 3G/4$.

Разумеется, вместо уравнения моментов (12.4) относительно точки A можно было бы написать уравнение моментов относительно точки B (или любой другой точки). При этом получилась бы система уравнений, эквивалентная использованной системе (12.3) и (12.4).

Сила натяжения веревки и силы взаимодействия в шарнире для рассматриваемой системы являются внутренними и поэтому не могут быть определены из условий равновесия всей системы в целом. Для определения этих сил не-

обходимо рассматривать условия равновесия отдельных частей системы. При этом удачным выбором точки, относительно которой составляется уравнение моментов сил, можно добиться упрощения алгебраической системы уравнений. Так, например, в данной системе можно рассмотреть условие равновесия моментов сил, действующих на левую половину лестницы, относительно точки C, в которой находится шарнир. При таком выборе точки силы, действующие в шарнире, не войдут в это условие, и мы сразу находим силу натяжения веревки T:

$$N_1 l \cos\alpha = Tl \sin\alpha, \qquad (12.5)$$

откуда, учитывая, что $N_1 = G/4$, получаем

$$T = \frac{G}{4}\operatorname{ctg}\alpha.$$

Условие (12.5) означает, что равнодействующая сил $\boldsymbol{T}$ и $\boldsymbol{N}_1$ проходит через точку C, т. е. направлена вдоль лестницы. Поэтому равновесие этой половинки лестницы возможно, только если сила $\boldsymbol{Q}_1$, действующая на нее в шарнире, также направлена вдоль лестницы (рис. 12.3, *в*), а ее величина равна величине равнодействующей сил $\boldsymbol{T}$ и $\boldsymbol{N}_1$:

$$Q_1 = \frac{G}{4\sin\alpha}.$$

Сила $\boldsymbol{Q}_2$, действующая в шарнире на другую половинку лестницы, на основании третьего закона Ньютона равна $\boldsymbol{Q}_1$ по величине и направлена противоположно. Направление силы $\boldsymbol{Q}_1$ можно было бы определить непосредственно из рис. 12.3, *в*, учитывая, что при равновесии тела под действием трех сил линии, по которым действуют эти силы, пересекаются в одной точке. Действительно, рассмотрим точку пересечения линий действия двух из этих трех сил и составим уравнение моментов относительно этой точки. Моменты первых двух сил относительно этой точки равны нулю; значит, должен равняться нулю и момент третьей силы, что возможно, только если линия ее действия также проходит через эту точку.

Иногда задачу статики можно решить, вообще не рассматривая условий равновесия, а используя закон сохранения энергии применительно к механизмам без трения: ни один механизм не дает выигрыша в работе. Этот закон

часто называют золотым правилом механики. Для иллюстрации такого подхода рассмотрим следующий пример: тяжелый груз весом G подвешен на невесомом шарнире с тремя звеньями (рис. 12.4). Какое натяжение должна выдержать нить, соединяющая точки A и B?

Попробуем с помощью этого механизма поднимать груз G. Отвязав нить в точке A, потянем ее вверх так, чтобы точка B медленно поднялась на расстояние Δh. Величина перемещения Δh ограничена тем, что сила натяжения нити T должна оставаться неизменной в процессе перемещения. Совершенная при этом работа $\Delta A = T\,\Delta h$. В результате груз G поднимается на высоту Δh_1, которая, как ясно из геометрических соображений, равна $3\Delta h$. Так как при отсутствии трения никаких потерь энергии не происходит, можно утверждать, что изменение потенциальной энергии груза, равное $G\,\Delta h_1$, определяется совершенной при подъеме работой. Поэтому

$$T = 3G.$$

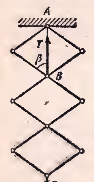


Рис. 12.4. В шарнире из трех звеньев точки A и B соединены нитью.

Очевидно, что для шарнира, содержащего произвольное число n одинаковых звеньев,

$$T = nG.$$

Нетрудно найти натяжение нити и в том случае, когда требуется учитывать вес самого шарнира G_1: совершаемую при подъеме работу следует приравнять сумме изменений потенциальных энергий груза и шарнира. Для шарнира из одинаковых звеньев центр тяжести его поднимается на $n\,\Delta h/2$. Поэтому

$$T = n\left(G + \frac{G_1}{2}\right).$$

Сформулированный принцип («золотое правило механики») применим и тогда, когда в процессе перемещений не происходит изменения потенциальной энергии, а механизм используется для преобразования силы. Редукторы, трансмиссии, вороты, системы рычагов и блоков — во всех таких системах величину преобразованной силы можно определить, приравнивая работы преобразованной и при-

ложенной сил. Другими словами, при отсутствии трения отношение этих сил определяется только геометрией устройства. Рассмотрим с этой точки зрения разобранный выше пример со стремянкой. Конечно, использовать стремянку в качестве подъемного механизма, т. е. поднимать человека, сближая половинки стремянки, вряд ли целесообразно. Однако это не может помешать нам применить описанный метод для нахождения силы натяжения веревки. Приравнивая работу, совершаемую при сближении частей стремянки, изменению потенциальной энергии человека на стремянке и связывая из геометрических соображений перемещение Δx нижнего конца лестницы с изменением высоты груза Δh (рис. 12.5), получаем, как и следовало ожидать, приведенный ранее результат:

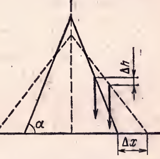


Рис. 12.5. Перемещение нижнего конца лестницы Δx и груза Δh при сближении половинок стремянки.

$$T = \frac{G}{4}\operatorname{ctg}\alpha.$$

Как мы уже отмечали, величину перемещения следует выбирать такой, чтобы в процессе перемещения можно было считать постоянной величину действующей силы. Легко убедиться, что в примере с шарниром это условие не накладывает ограничений на величину перемещения, так как сила натяжения нити не зависит от угла β (рис. 12.4). Напротив, в задаче о стремянке перемещение следует выбирать малым, ибо сила натяжения веревки зависит от α.

Равновесие бывает устойчивым, неустойчивым и безразличным. Равновесие устойчиво, если при малых перемещениях тела из положения равновесия действующие силы стремятся вернуть его обратно, и неустойчиво, если силы уводят его дальше от положения равновесия. Если же при малых смещениях действующие на тело силы и их моменты по-прежнему уравновешиваются, то равновесие безразличное.

Устойчивому равновесию соответствует минимум потенциальной энергии тела по отношению к ее значениям в соседних положениях тела. Этим свойством часто удобно пользоваться при отыскании положения равновесия и исследовании характера равновесия.

Рассмотрим пример исследования устойчивости равновесия. Пусть имеются два круглых карандаша радиусами R

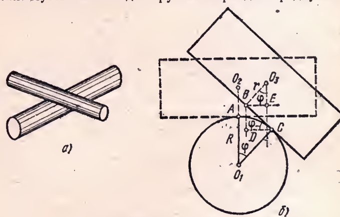


Рис. 12.6. Равновесие круглого карандаша.

и r. Один из них расположен горизонтально, другой уравновешен на нем в горизонтальном положении так, что оси карандашей взаимно перпендикулярны (рис. 12.6, *а*). При каком соотношении между радиусами равновесие устойчиво? На какой максимальный угол можно при этом отклонить от горизонтали верхний карандаш? Коэффициент трения равен μ.

На первый взгляд может показаться, что равновесие верхнего карандаша вообще неустойчиво, так как центр тяжести верхнего карандаша лежит выше оси, вокруг которой он может поворачиваться. Однако здесь положение оси вращения не остается неизменным, поэтому этот случай требует специального исследования.

Поскольку верхний карандаш уравновешен в горизонтальном положении, центры тяжести карандашей O_1 и O_2 лежат на одной вертикали (рис. 12.6, *б*).

Отклоним верхний карандаш на некоторый угол φ от горизонтали. При отсутствии трения покоя он немедленно

соскользнул бы вниз. Чтобы не думать пока о возможном соскальзывании, будем считать трение достаточно большим. При этом верхний карандаш «прокатывается» по нижнему без проскальзывания. Точка опоры из положения A перемещается в новое положение C, а та точка, которой верхний карандаш до отклонения опирался о нижний, переходит в положение B. Поскольку проскальзывание отсутствует, длина дуги AC равна длине отрезка BC: $\smile AC = R\varphi = BC$. Центр тяжести верхнего карандаша O_2 переходит в положение O_3. Если вертикаль, проведенная через O_3, проходит левее новой точки опоры C, то сила тяжести стремится вернуть верхний карандаш в положение равновесия. Выразим это условие математически. Проведя вертикаль через точку B, видим, что должно быть выполнено условие

$$BE < DC. \qquad (12.6)$$

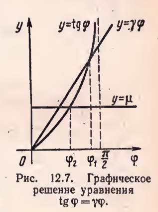


Рис. 12.7. Графическое решение уравнения $\operatorname{tg}\varphi = \gamma\varphi$.

Так как $BE = r\sin\varphi$, а $DC = BC\cos\varphi = R\varphi\cos\varphi$, то из условия (12.6) получаем

$$\frac{r}{R} < \frac{\varphi}{\operatorname{tg}\varphi}. \qquad (12.7)$$

Поскольку $\operatorname{tg}\varphi > \varphi$ $(0 < \varphi < \pi/2)$, сила тяжести будет стремиться возвратить верхний карандаш в положение равновесия только при $r/R < 1$.

Следовательно, устойчивое равновесие верхнего карандаша на нижнем возможно только тогда, когда его радиус меньше радиуса нижнего карандаша.

Для ответа на второй вопрос следует выяснить, какие причины ограничивают величину допустимого угла отклонения. Во-первых, при больших углах отклонения вертикаль, проведенная через центр тяжести верхнего карандаша, может пройти правее точки опоры C. Из условия (12.7) видно, что при заданном отношении радиусов карандашей $\gamma = R/r$ максимальный угол отклонения φ_1 определяется уравнением $\operatorname{tg}\varphi_1 = \gamma\varphi_1$. Решение этого трансцендентного уравнения легко найти графически (рис. 12.7). Во-вторых,

максимальное значение угла отклонения ограничивается величиной трения: карандаш не должен соскальзывать, т. е. для предельного угла φ_2 получаем $\operatorname{tg}\varphi_2=\mu$ (вспомните условие равновесия на наклонной плоскости). Решение этого уравнения также показано на рис. 12.7. Очевидно, что максимально допустимый угол отклонения равен меньшему из φ_1 и φ_2. Поскольку $\gamma>1$, а коэффициент трения μ обычно меньше единицы, то максимально допустимый угол отклонения практически всегда определяется условием соскальзывания, т. е. углом φ_2.

ДВИЖЕНИЕ ЖИДКОСТЕЙ И ГАЗОВ

§ 13. Гидростатика. Плавание тел

Главное макроскопическое проявление различия между твердыми телами, с одной стороны, и жидкостями и газами, с другой, связано с их поведением при изменении формы. Твердые тела деформируются только под действием сил. Медленное изменение формы жидкости без изменения ее объема может происходить под действием сколь угодно малой силы. В предельном случае бесконечно медленного изменения формы жидкости вообще никаких сил не требуется. Жидкости и газы ведут себя как упругие тела только в отношении изменения их объема.

Так как изменение формы выделенного объема всегда связано с касательными напряжениями в среде, то при равновесии касательные напряжения в жидкостях и газах не возникают. Отсюда следует, что в состоянии равновесия величина нормального напряжения — давления — не зависит от ориентации площадки, на которую оно действует.

Для доказательства этого утверждения рассмотрим мысленно выделенный внутри жидкости объем в виде прямой треугольной призмы (рис. 13.1, *а*). На боковые грани этой призмы со стороны окружающей жидкости действуют силы $\boldsymbol{F}_1$, $\boldsymbol{F}_2$, $\boldsymbol{F}_3$, направленные перпендикулярно к соответствующим граням. В равновесии векторная сумма этих сил равна нулю и, следовательно, векторы этих сил образуют треугольник (рис. 13.1, *б*). Этот треугольник подобен треугольнику, лежащему в основании призмы (как треугольники со взаимно перпендикулярными сторонами).

Из подобия треугольников следует, что отношение величины силы F_i к площади грани S_i, на которую она действует, одинаково для всех граней. Но отношение F_i/S_i представляет собой давление в жидкости, что и доказывает

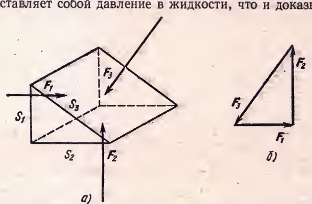


Рис. 13.1. Равновесие мысленно выделенного объема жидкости под действием сил давления (к доказательству закона Паскаля).

независимость давления от ориентации площадки в жидкости или газе.

Экспериментально этот закон был открыт Паскалем и носит его имя.

Не оказывая сопротивления изменению формы, жидкости и газы тем не менее сопротивляются изменению объема. Газы обладают способностью к неограниченному расширению, т. е. заполняют полностью предоставленный им объем. Напротив, для жидкости характерен определенный собственный объем, который лишь незначительно меняется при изменении внешнего давления. Во многих случаях изменение объема жидкости бывает столь малым, что им можно полностью пренебречь и рассматривать жидкость как несжимаемую, т. е. имеющую постоянную плотность. Такая жидкость называется абсолютно несжимаемой.

Как и в случае абсолютно твердого тела, применимость представления об абсолютно несжимаемой жидкости определяется не столько свойствами самой жидкости, сколько условиями, в которых она находится. Например, при изучении распространения звуковых волн в жидкости всегда необходимо учитывать ее сжимаемость, в то время

как при изучении движения потоков не только жидкость, но и газ часто можно рассматривать как несжимаемые.

В поле земного тяготения давление в жидкости, хотя и не зависит от ориентации площадки, оказывается разным на разной глубине. Для несжимаемой жидкости зависимость давления от глубины может быть легко найдена. Рассматривая условия равновесия мысленно выделенного объема жидкости в виде вертикально расположенного цилиндра (рис. 13.2) и учитывая наряду с силами давления со стороны остальной жидкости также и силу тяжести $G=\rho ghS$, действующую на выделенный объем, получим

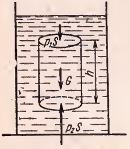


Рис. 13.2. Равновесие мысленно выделенного объема жидкости при наличии силы тяжести.

$$p_2-p_1=\rho gh, \qquad (13.1)$$

где p_1 и p_2 — давления в жидкости у верхнего и нижнего оснований цилиндра, h — высота цилиндра, ρ — плотность жидкости. Обусловленное силой тяжести давление в жидкости носит название гидростатического. Хотя гидростатическое давление зависит от глубины h, оно не зависит от ориентации малой площадки, на которую оно действует. Приведенное выше доказательство закона Паскаля остается

Рис. 13.3. Силы давления жидкости на дно сосудов одинаковы во всех трех случаях.

справедливым и в этом случае, так как размер призмы, использовавшейся при доказательстве закона (рис. 13.1, *а*), может быть выбран сколь угодно малым.

Гидростатическое давление жидкости на дно сосуда определяется высотой столба и плотностью жидкости и не зависит от формы стенок. Поэтому сила давления жидкости на дно сосудов, изображенных на рис. 13.3, у которых площади дна равны, одинакова, несмотря на то, что сосуды

содержат разное количество жидкости. Может показаться, что при взвешивании этих сосудов весы покажут один и тот же вес, так как показания весов определяются силой, с которой дно сосуда действует на чашку весов. Это, разумеется, не так, и проще всего в этом убедиться, сообразив, что показания весов не могут зависеть от того, поставлен ли сосуд дном на чашку весов или подвешен к ней за горлышко.

Наличие обусловленного полем тяжести гидростатического давления приводит к существованию статической подъемной силы, действующей на погруженные в жидкость тела. Закон, определяющий величину выталкивающей силы, был открыт великим Архимедом: эта сила равна весу жидкости, объем которой совпадает с объемом погруженной в жидкость части тела. В справедливости этого утверждения легко убедиться следующим образом. Выделим в жидкости объем произвольной формы. В состоянии равновесия действующая на жидкость в выделенном объеме сила тяжести уравновешивается силами гидростатического давления, действующими на поверхность выделенного объема со стороны окружающей жидкости. Если заменить выделенный объем жидкости твердым телом точно такой же формы, то действующие на поверхность тела силы гидростатического давления окружающей жидкости, очевидно, не изменятся и, следовательно, их равнодействующая будет по-прежнему равна весу выделенного объема жидкости и направлена вертикально вверх. Это и есть архимедова выталкивающая сила, действующая на погруженное в жидкость тело. Очевидно, что линия действия этой силы проходит через центр масс выделенного объема жидкости и не зависит от того, где расположен центр масс погруженного тела.

Если средняя плотность тела меньше плотности жидкости, то часть тела выступает над поверхностью. В этом случае говорят, что тело плавает. Для строительства кораблей большое значение имеет вопрос устойчивости плавания. Эта устойчивость определяется положением метацентра. На рис. 13.4 изображен корабль, накрененный на некоторый угол α от вертикального положения. При этом центр тяжести вытесненной им в этом наклонном положении воды находится в точке B, смещенной из плоскости симметрии корабля NN в ту же сторону, куда накренился корабль. Проведем через точку B вертикаль, которая

представляет собой линию действия выталкивающей силы. Точка C пересечения этой линии с плоскостью симметрии корабля называется метацентром. Если метацентр лежит выше центра тяжести корабля O, то момент выталкивающей силы относительно центра тяжести корабля стремится возвратить корабль в вертикальное положение, т. е. корабль плавает устойчиво. Если же метацентр лежит ниже центра тяжести корабля, то плавание корабля в вертикальном положении будет неустойчивым.

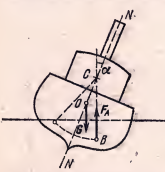


Рис. 13.4. К устойчивости корабля в вертикальном положении.

Если средняя плотность тела равна плотности жидкости, то архимедова выталкивающая сила равна весу тела. В этом случае тело целиком погружено в жидкость и находится в состоянии равновесия. Если бы тело, как и жидкость, было абсолютно несжимаемым (или сжимаемости тела и жидкости были одинаковы), то это равновесие было бы безразличным. Но у реальных твердых материалов сжимаемость, как правило, меньше сжимаемости жидкости. Тела из таких материалов при равенстве их плотности плотности жидкости должны были бы устойчиво плавать в погруженном состоянии на некоторой глубине. Но практически так никогда не бывает, так как совпадение плотности жидкости и плотности материала почти невероятно.

Однако сделать *среднюю* плотность твердого тела равной плотности жидкости не представляет труда. Средняя плотность подводной лодки в погруженном состоянии как раз равна плотности воды. Может ли такая лодка зависнуть на некоторой глубине в погруженном состоянии? Оказывается, нет. Пусть на некоторой глубине средняя плотность лодки равна плотности воды. Представим себе, что вследствие случайных причин лодка погрузилась чуть глубже. Сжимаемость лодки определяется не столько сжимаемостью материала, из которого она сделана, сколько жесткостью ее конструкции. Практически эта сжимаемость всегда гораздо больше сжимаемости воды. Поэтому при небольшом

погружеиии увелнчнвшееся гидростатическое давленне прнведет к деформацнн корпуса, н средияя плотность лодкн станет больше плотностн воды — лодка будет погружаться еще глубже. Совершеино аиалогично при случайном уменьшеиин глубнны погружения условие равновесия также нарушнтся, н лодка будет всплывать.

Можио проделать простой опыт, иллюстрнрующнй условия плавання тела в погружеином состояннн. В высоком цнлиндрическом сосуде с водой плавает перевериутая отверстнем вниз пробирка, частичио заполнениая воздухом (рис. 13.5). Колнчество воздуха в пробирке иужио подобрать такнм образом, чтобы из воды чуть высовывалось только донышко пробнркн: средняя плотность плавающей пробирки с воздухом иемиого меньше плотности воды. Отверстне цилнндрнческого сосуда затягнвается тонкой прочиой резиновой пленкой. При иажатии пальцем иа пленку давлеиие воздуха над поверхностью воды в цнлиндре возрастает. В результате воздух в пробнрке сжимается, и средияя плотность пробнрки с воздухом становится больше плотности воды — пробнрка тоиет. Если отпустить пленку, то давлеине примет первоиачальиое зиачеиие, и пробирка всплывет.

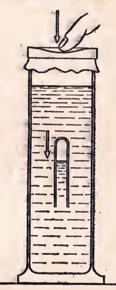


Рис. 13.5. Демонстрация плавания тела в погруженном состоянии.

Разумеется, когда пробирка иаходнтся на иекоторой глубиие, можио, измеияя иажатие иа плеику, добнться равенства плотиости воды и средней плотиостн пробирки. Но равновесие пробирки, как н в случае подводной лодки, будет неустойчивым. Добнться того, чтобы пробнрка оставалась неподвижиой иа некоторой глубине, можио только дниамически, периодически увелнчивая н умеиьшая иажатне на плеику. При этом средияя плотиость пробиркн будет то больше, то меиьше плотности воды. Но благодаря иериции пробнркн и вязкостн жндкости можно добнться того, что колебаиия пробнрки будут практически иезаметиыми.

В отличие от рассмотренных выше примеров, дирижабль с жесткой оболочкой может устойчиво висеть в воздухе. Пусть на некоторой высоте средняя плотность дирижабля равна плотности воздуха на этой высоте. Вследствие жесткости оболочки дирижабля можно считать, что его средняя плотность при изменении внешнего давления остается неизменной. Поэтому при случайном уменьшении высоты подъемная сила возрастает, так как плотность воздуха при этом увеличивается.

§ 14. Движение идеальной жидкости

При кинематическом описании движения жидкости или газа можно поступать следующим образом. Будем следить за определенной точкой пространства и фиксировать величину и направление скоростей различных частиц жидкости, которые в разные моменты времени проходят через

Рис. 14.1. Скорости частиц движущейся жидкости и линии тока.

эту точку. Если проделать это для всех точек пространства и указать скорости частиц жидкости во всех точках в определенный момент времени, то получится мгновенная картина распределения скоростей в движущейся жидкости — так называемое поле скоростей. Линии, касательные к которым во всех точках совпадают с направлениями скоростей жидкости в этих точках, называются линиями тока (рис. 14.1).

При стационарном течении жидкости поле скоростей, а следовательно, и линии тока не меняются со временем. В этом случае линии тока совпадают с траекториями от-

дельных частиц жидкости, так как каждая частица жидкости приходит в данную точку с той же самой скоростью.

Часть жидкости, ограниченная линиями тока, называется трубкой тока (рис. 14.2). Такая мысленно выделенная в потоке часть жидкости — трубка тока,— подобно жидкости в настоящей трубе, движется, нигде не пересекая боковой поверхности трубки. При стационарном течении количество жидкости, пересекающей в единицу времени сечение S_1, т. е. «втекающей» в выделенную часть трубки, равно количеству жидкости, «вытекающей» через сечение S_2. Если выбрать трубку тока с поперечным сечением ΔS настолько малым, чтобы скорость жидкости во всех точках сечения была одинаковой, причем это сечение ориентировано перпендикулярно линиям тока, то количество жидкости Δm, протекающей через это сечение за время t, будет равно

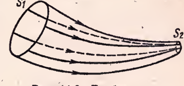


Рис. 14.2. Трубка тока.

$$\Delta m = \rho v \, \Delta S \, t. \qquad (14.1)$$

В стационарном потоке величина Δm одна и та же для любого поперечного сечения выбранной трубки тока, поэтому согласно (14.1)

$$\rho_1 v_1 \, \Delta S_1 = \rho_2 v_2 \, \Delta S_2. \qquad (14.2)$$

Если жидкость можно рассматривать как несжимаемую, то $\rho_1 = \rho_2$ и условие (14.2) принимает вид

$$v_1 \, \Delta S_1 = v_2 \, \Delta S_2. \qquad (14.3)$$

Это соотношение называется уравнением неразрывности. Полученный результат (14.3) справедлив для выбранной трубки тока. При изучении движения потоков жидкости на такие трубки можно разбить все пространство, занимаемое жидкостью.

Динамика движения реальной жидкости очень сложна. Для упрощения ее описания в некоторых случаях можно пренебречь силами внутреннего трения. Такую жидкость называют идеальной. При движении идеальной жидкости не происходит превращения механической энергии во

внутреннюю, т. е. механическая энергия жидкости сохраняется. Закон сохранения механической энергии для идеальной несжимаемой жидкости выражается уравнением Бернулли.

Рассмотрим часть жидкости, заключенную между сечениями ΔS_1 и ΔS_2 некоторой трубки тока, расположенными на высотах h_1 и h_2 соответственно (рис. 14.3). За промежуток времени Δt эта жидкость смещается вдоль трубки тока и занимает новое положение между сечениями $\Delta S_1'$ и $\Delta S_2'$.

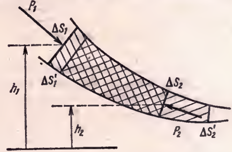


Рис. 14.3. К выводу уравнения Бернулли.

Для малого промежутка времени Δt можно пренебречь различием между площадями ΔS и $\Delta S'$ старых и новых сечений и различием в их высотах. Подсчитаем работу, совершаемую внешними силами над выделенной жидкостью за время Δt. Силы давления, действующие на боковую поверхность трубки тока, работы не совершают, так как действуют перпендикулярно перемещению. Работа силы давления в сечении ΔS_1 равна $p_1\Delta S_1 v_1 \Delta t$, работа в сечении ΔS_2 равна $-p_2\Delta S_2 v_2 \Delta t$, так что полная работа внешних сил

$$\Delta A = p_1\,\Delta S_1\,v_1\,\Delta t - p_2\,\Delta S_2\,v_2\,\Delta t. \qquad (14.4)$$

В силу стационарности движения энергия жидкости между сечениями $\Delta S_1'$ и ΔS_2 не меняется. Эта часть жидкости показана на рис. 14.3 двойной штриховкой. Поэтому изменение энергии рассматриваемой жидкости равно энергии части жидкости между сечениями ΔS_2 и $\Delta S_2'$ минус энергия части жидкости между сечениями ΔS_1 и $\Delta S_1'$. Потенциальная энергия части жидкости между ΔS_2 и $\Delta S_2'$ равна $\rho\,\Delta S_2 v_2\,\Delta t g h_2$, ее кинетическая энергия равна $^1/_2\rho\,\Delta S_2 v_2\,\Delta t v_2^2$. Аналогично записываются выражения для

энергии жидкости между сечениями ΔS_1 и $\Delta S_1'$. Поэтому изменение энергии всей выделенной части жидкости в рассматриваемой трубке тока за время Δt равно

$$\Delta E = \rho\,\Delta S_2\,v_2\,\Delta t\,gh_2 + \frac{1}{2}\rho\,\Delta S_2\,v_2\,\Delta t\,v_2^2 - \left(\rho\,\Delta S_1\,v_1\,\Delta t\,gh_1 + \frac{1}{2}\rho\,\Delta S_1\,v_1\,\Delta t\,v_1^2\right). \quad (14.5)$$

На основании закона сохранения механической энергии работа внешних сил (14.4) равна изменению энергии системы (14.5). Учитывая уравнение неразрывности (14.3), получим

$$p_1 + \rho gh_1 + \frac{1}{2}\rho v_1^2 = p_2 + \rho gh_2 + \frac{1}{2}\rho v_2^2. \quad (14.6)$$

Это и есть уравнение Бернулли. Оно было выведено для достаточно узкой трубки тока и, строго говоря, справедливо, когда эта трубка сжимается в линию тока. Поэтому

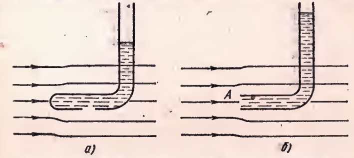


Рис. 14.4. Манометрическая трубка в потоке жидкости.

сумма $p + \rho gh + \frac{1}{2}\rho v^2$ остается неизменной вдоль одной и той же линии тока.

В неподвижной жидкости в состоянии равновесия согласно закону Паскаля давление не зависит от ориентации площадки. А как обстоит дело в движущейся жидкости? Уравнение Бернулли дает возможность ответить на этот вопрос в случае стационарного течения идеальной несжимаемой жидкости. Оказывается, что величина измеряемого неподвижным манометром давления зависит от ориентации площадки в потоке.

Представим себе манометр в виде изогнутой трубки, передняя часть которой, обращенная навстречу потоку, запаяна, а в боковой стенке имеется параллельное скорости обтекающей жидкости отверстие (рис. 14.4, *а*). Такая трубка искажает поток только вблизи ее переднего конца, а вблизи отверстия поток практически не меняется. Поэтому давление здесь такое же, как и во всех других точках линии тока, проходящей вблизи отверстия. Соединенный с такой трубкой манометр измеряет давление жидкости p, входящее в уравнение Бернулли. Такое же давление покажет произвольно ориентированный манометр, движущийся вместе с потоком.

Если же взять трубку с открытым передним концом, обращенным навстречу потоку жидкости (рис. 14.4, *б*), то показание соединенного с ней манометра будет больше p. Поясним это. Линии тока вблизи такой трубки показаны на рис. 14.4, *б*. Так как жидкость внутри трубки неподвижна, то линия тока, упирающаяся в открытый конец трубки, обрывается в точке A и скорость жидкости в этой точке обращается в нуль. Обозначим давление в этой точке через p_1, а давление и скорость в потоке вдали от трубки через p и v.

Применяя к выделенной линии тока уравнение Бернулли, получим

$$p_1 = p + \frac{1}{2}\rho v^2. \tag{14.7}$$

Именно это давление и показывает соединенный с трубкой манометр. По измерениям величин p и p_1, т. е. располагая трубками обоих типов, можно рассчитать скорость потока v.

С помощью уравнения Бернулли легко оценить скорость истечения жидкости v из шприца. Будем считать жидкость идеальной. Пусть на поршень шприца, который имеет площадь S_0, действует внешняя сила F (рис. 14.5) и струя жидкости вытекает из иглы с отверстием площади S. Рассмотрим линию тока, проходящую вдоль оси симметрии шприца, и применим к ней уравнение Бернулли. Обозначая скорость поршня и, следовательно, жидкости вблизи него через v_0, имеем

$$\frac{F}{S_0} + \frac{1}{2}\rho v_0^2 = \frac{1}{2}\rho v^2. \tag{14.8}$$

Из уравнения неразрывности (14.3) вытекает, что $S_0 v_0 = Sv$. Выражая отсюда v_0 и подставляя в (14.8), получим

$$\frac{F}{S_0} = \frac{1}{2}\rho v^2 \left[1 - \left(\frac{S}{S_0}\right)^2\right]. \qquad (14.9)$$

Обычно площадь отверстия иглы во много раз меньше площади поршня шприца: $S \ll S_0$. Тогда, пренебрегая квадратом

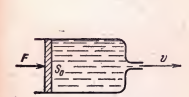


Рис. 14.5. К оценке скорости истечения жидкости из шприца.

Рис. 14.6. Истечение жидкости из отверстия в стенке сосуда.

отношения S/S_0 по сравнению с единицей, найдем скорость истечения:

$$v = \sqrt{\frac{2F}{\rho S_0}}.$$

Как вытекает налитая в широкий сосуд жидкость из небольшого отверстия в дне или боковой стенке под действием силы тяжести (рис. 14.6)? Скорость истечения идеальной несжимаемой жидкости легко найти с помощью уравнения Бернулли. Рассмотрим линию тока, начинающуюся вблизи свободной поверхности жидкости и проходящую вдоль оси отверстия. Поскольку скорость жидкости вблизи поверхности в широком сосуде пренебрежимо мала, то уравнение Бернулли имеет вид

$$\rho g h = \frac{1}{2}\rho v^2,$$

откуда

$$v = \sqrt{2gh}. \qquad (14.10)$$

Таким образом, скорость истечения идеальной жидкости из отверстия в сосуде такая же, как и при свободном

падении тела с высоты h. Этот факт был впервые установлен Торричелли.

Более сложным является вопрос о форме струи вытекающей жидкости. Оказывается, что форма струи зависит от устройства отверстия. Сравнительно просто исследовать предельные случаи, показанные на рис. 14.7, *а* и *б*. В случае *а* линии тока в отверстии перед истечением постепенно

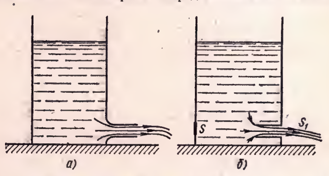


Рис. 14.7. Зависимость сечения струи жидкости от устройства отверстия.

меняют направление на параллельное оси трубки. В результате площадь сечения вытекающей струи равна площади сечения отверстия трубки и сжатия струи не происходит. В случае *б* частицы жидкости вблизи отверстия имеют скорости в поперечных направлениях, что приводит к сжатию струи. Величину сжатия для этого случая можно рассчитать с помощью закона сохранения импульса. Будем рассуждать следующим образом. Всюду вблизи боковых стенок сосуда скорость движения жидкости пренебрежимо мала и давление равно гидростатическому. Силы давления жидкости на стенки сосуда взаимно уравновешиваются всюду, за исключением участка, лежащего точно напротив отверстия и имеющего ту же площадь S, что и отверстие. Импульс этой неуравновешенной силы за время Δt равен $\rho g h S\, \Delta t$. На основании закона сохранения импульса точно такой же по величине импульс должна унести вытекающая за это время Δt жидкость. Этот импульс равен произведению массы вытекающей жидкости на скорость ее истечения v. Если площадь сечения струи после

сжатия S_1, то импульс жидкости равен $\rho\ S_1 v\ \Delta t\ v$. Поэтому

$$\rho g h S\,\Delta t = \rho S_1 v^2\,\Delta t.$$

Подставляя сюда скорость истечения жидкости (14.10), получим $S_1 = S/2$: поперечное сечение вытекающей струи оказывается вдвое меньше площади отверстия. При всех других формах отверстий, отличающихся от изображенных

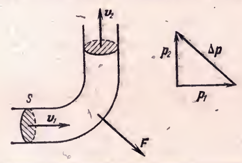


Рис. 14.8. Реакция струи жидкости при течении по изогнутой трубе.

на рис. 14.7, сжатие струи заключено в промежутке между этими предельными случаями.

Закон сохранения импульса позволяет объяснить реакцию струи жидкости, которая течет по изогнутой трубе постоянного сечения площади S (рис. 14.8). При стационарном течении импульс любого элемента жидкости изменяется только по направлению, оставаясь неизменным по величине. В трубе, изогнутой под прямым углом, изменение импульса жидкости за время Δt, как видно из рис. 14.8, равно

$$\Delta \mathbf{p} = \mathbf{p}_2 - \mathbf{p}_1 = \rho S v\,(\mathbf{v}_2 - \mathbf{v}_1)\,\Delta t, \qquad (14.11)$$

где $\mathbf{v}_1$ и $\mathbf{v}_2$ — равные по величине скорости жидкости до и после изгиба трубы: $v_1 = v_2 = v$. Таким образом, действующая на трубу сила, обусловленная движением жидкости, равна

$$F = \sqrt{2}\,\rho S v^2. \qquad (14.12)$$

Направление этой силы указано на рисунке. Разобранный пример объясняет принцип действия гидравлических турбин.

В заключение рассмотрим явление так называемого гидравлического удара. Нередко можно видеть, как в

твердых камнях выбиты углубления в тех местах, куда попадают отдельные падающие сверху капли воды. Дело в том, что при ударе капель о преграду, в отличие от постоянно действующей струи, происходит внезапно возникающий контакт струи с преградой. В непрерывной струе, как мы видели, на поставленную поперек потока площадку действует добавочная сила $\rho v^2/2$ на единицу площади. Если же неподвижная площадка появляется в потоке внезапно, то набегающая на нее жидкость вынуждена затормозиться. Абсолютно несжимаемая жидкость, движущаяся по трубе, при мгновенном перекрывании трубы остановилась бы вся сразу, что привело бы к бесконечно большой силе давления на преграду. Поэтому представление об абсолютно несжимаемой жидкости в таких условиях неприменимо. В сжимаемой жидкости при внезапном появлении преграды за время Δt остановится только та часть жидкости, до которой успеет дойти волна деформации, распространяющаяся в жидкости навстречу потоку от преграды. Такая волна распространяется со скоростью, равной скорости звука c в данной жидкости. Закон сохранения импульса позволяет рассчитать величину силы F, действующей на внезапно возникающую в трубе сечением S перегородку. Пусть до появления преграды жидкость в трубе имела скорость v. Учитывая, что масса останавливающейся за время Δt жидкости равна $\rho S c\,\Delta t$, имеем

$$F\,\Delta t = \rho S c\,\Delta t\, v,$$

откуда для развивающегося при гидравлическом ударе добавочного давления $p = F/S$ имеем

$$p = \rho c v. \qquad (14.13)$$

Скорость звука в воде равна примерно 1500 м/с. Поэтому в потоке, имеющем скорость 10 м/с, давление (14.13), развиваемое при гидравлическом ударе, как нетрудно убедиться, в 300 раз больше давления $\rho v^2/2$ постоянно действующей струи воды.

Явления, связанные с гидравлическим ударом, весьма разнообразны. Например, во время шторма на море можно наблюдать, как волны, бьющие в вертикальную стенку набережной, образуют всплески, имеющие огромную высоту, в десятки раз превосходящую высоту волн на море.

§ 15. Вязкая жидкость. Обтекание тел

В целом ряде практнческн важиых случаев поведеиие обычной жндкости в пределах известной точностн экспернмеита согласуется с результатами, предсказываемымн теорией движения идеальиой несжимаемой жндкости. Но есть немало примеров, когда иельзя пренебрегать виутреиинм треиием в жидкости, называемым вязкостью. Большннство интересиых явлеинй прн двнжеиии жидкости так илн ниаче связано нменно с этим свойством, а в иекоторых случаях вязкость жидкости является определяющей.

При изучеиин идеальиой жидкости выдвигалось требованне, чтобы иормальиая составляющая скоростн жндкостн на ее граинце с твердым телом обращалась в нуль. Касательиая составляющая скорости могла нметь пронзвольное значенне. Но оказывается, что во всех случаях, где это можно проверить эксперимеитально, скорость жидкости иа поверхностн твердого тела строго равиа нулю, т. е. жидкость как бы прнлнпает к поверхности, которую она обтекает. Вблнзи поверхностн твердого тела возннкает так иазываемый пограиичный слой жидкостн, в котором скорость нарастает от нуля до значення скорости в потоке. Влияине вязкости иа движение остальиой частн жндкости мало. Но еслн, напрнмер, вязкая жидкость двнжется по достаточно узкой трубе, то пограничный слой может заполннть весь объем текущей жндкости, и прн изученнн этого двнжения преиебрегать вязкостью нельзя. Такое течение имеет очеиь мало общего с двнженнем идеальной жидкости.

Как уже отмечалось, в статнческом случае, т. е. в равновесни, ннкаких напряженнй сдвига в жндкости иет. Однако в двнжущейся жндкости такне напряжения могут быть. Вязкость как раз н опнсывает такне снлы, возннкающне в двнжущейся жидкости. В отличне от твердых тел, где сдвиговое касательное иапряженне определяется величнной деформацни сдвнга, в жидкости такое напряжение определяется скоростью деформации. Другнми словамн, жидкостн оказывают вязкое, а не упругое сопротивленне прн измененни формы.

Для того чтобы ввести количественную характернстнку вязкости жидкости, рассмотрим следующий опыт. Пусть жидкость находнтся между двумя твердымн плоскимн

параллельными пластинами (рис. 15.1). Нижняя пластина неподвижна, а верхняя движется параллельно нижней с малой скоростью v. Опыт показывает, что для поддержания равномерного движения верхней пластины необходима сила F, направленная вдоль пластины и пропорциональная площади пластины S, скорости v и обратно пропорциональная расстоянию d между пластинами:

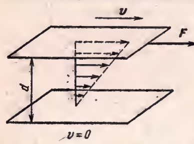


Рис. 15.1. Вязкая жидкость между плоскими пластинами.

$$F=\eta\frac{Sv}{d}. \qquad (15.1)$$

Подчеркнем, что благодаря «прилипанию» жидкости к поверхности пластины эта сила характеризует внутреннее трение, т. е. трение между проскальзывающими друг относительно друга слоями жидкости, а не между жидкостью и твердым телом.

Величина η в (15.1) описывает вязкие свойства жидкости и называется коэффициентом вязкости. Вязкость жидкости сильно зависит от ее температуры. Так, например, вязкость воды при повышении ее температуры от 0 до 20° С уменьшается почти вдвое.

При наличии вязкости, т. е. сил внутреннего трения, тормозящих движение жидкости, для поддержания стационарного течения в горизонтальной трубе неизменного сечения необходимо поддерживать постоянную разность давлений на концах трубы. Напомним, что в идеальной жидкости при таком движении давление, как это следует из уравнения Бернулли, одинаково вдоль всей трубы.

Течение жидкости в цилиндрической трубе, при котором скорости частиц жидкости всюду направлены вдоль оси трубы, называется ламинарным или слоистым. Такое течение возможно только при не очень большой скорости потока вязкой жидкости в трубах малого поперечного сечения. С увеличением скорости или с увеличением площади сечения трубы характер течения принципиально изменяется. Вместо слоистого течения возникает носящее нерегулярный характер завихренное, или турбулентное, течение. Изменение характера течения можно наблюдать

в эксперименте со стеклянными трубками различного сечения при различных перепадах давления, т. е. при различных скоростях жидкости. Линии тока при стационарном течении можно сделать видимыми, впуская во входное сечение стеклянной трубки окрашенную струйку жидкости. При небольшой скорости потока в узкой трубке подкрашенная струйка движется ровно и параллельно оси трубки.

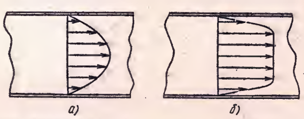


Рис. 15.2. Профиль скоростей при ламинарном (*а*) и турбулентном (*б*) течении жидкости по трубе.

Затем при постепенном увеличении скорости потока внезапно начинается нерегулярное движение, которое постепенно захватывает всю трубку,— струйка, ровная у входа, разбивается на множество извилистых струек. Такие нерегулярные изменения движения происходят не из-за изменения внешних условий, а вследствие неустойчивости ламинарного течения при больших скоростях.

При стационарном турбулентном движении скорость жидкости в данном месте не остается постоянной, а совершает хаотические колебания и по величине и по направлению. Но средняя скорость в данном месте трубы будет постоянна и направлена вдоль оси трубы. На рис. 15.2, *а* показано распределение скорости жидкости по сечению трубы при ламинарном течении, а на рис. 15.2, *б* — распределение средней скорости при установившемся турбулентном течении. В турбулентном потоке, как видно из рисунка, можно четко выделить пограничный слой жидкости вблизи стенок трубы, где средняя скорость быстро спадает до нуля, в то время как при ламинарном течении такого четкого пограничного слоя нет, так как скорость изменяется за счет вязкости по всему сечению трубы. Другими словами, в этом случае вся труба находится в пределах пограничного слоя.

Неустойчивость ламинарного течения и возникновение турбулентности — сложные вопросы, до конца не выясненные и в настоящее время.

Большое практическое значение имеет вопрос о силах, действующих на твердое тело, движущееся в неподвижной жидкости или газе. Часто оказывается более удобным рассматривать обтекание неподвижного тела набегающим потоком жидкости. Оба подхода эквивалентны в силу механического принципа относительности.

Разложим полную силу $\boldsymbol{F}$, действующую на тело со стороны потока, на две составляющие: в направлении потока $\boldsymbol{F}_{\parallel}$ и перпендикулярную потоку $\boldsymbol{F}_{\perp}$. Силу $\boldsymbol{F}_{\parallel}$ по общепринятой терминологии называют лобовым сопротивлением, силу $\boldsymbol{F}_{\perp}$ — подъемной силой.

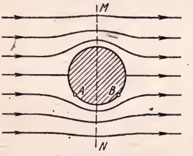


Рис. 15.3. Обтекание симметричного тела потоком идеальной жидкости.

При стационарном обтекании твердого тела потоком идеальной жидкости лобовое сопротивление должно вообще отсутствовать. Чтобы убедиться в этом, рассмотрим обтекание симметричного тела, изображенного на рис. 15.3. Линии тока симметричны относительно плоскости MN, а скорости частиц жидкости в соответствующих точках перед и за телом равны по величине и отличаются только по направлению. Давления в этих точках одинаковы согласно уравнению Бернулли. Теперь легко сообразить, что составляющие сил давления в точках A и B, направленные параллельно потоку, компенсируют друг друга. Так как полная сила, действующая на тело со стороны потока, равна сумме сил давления жидкости, действующих на отдельные элементы поверхности тела, то лобовое сопротивление отсутствует.

В рассмотренном на рис. 15.3 случае в силу симметрии картины обтекания тела равно нулю не только лобовое сопротивление, но и подъемная сила. Можно показать, что этот результат — равенство нулю полной силы, действующей на тело со стороны потока идеальной жидкости,—

справедлив для тела произвольной формы, имеющего конечные размеры. В этом состоит так называемый парадокс Даламбера.

Иначе обстоит дело при обтекании тела, имеющего бесконечные размеры в направлении, перпендикулярном потоку. Лобовое сопротивление в идеальной жидкости отсутствует и в этом случае, в то время как подъемная сила может быть отлична от нуля. Теория подъемной силы для крыла бесконечного размаха была создана Н. Е. Жуковским, который показал, что для возникновения подъемной силы необходимо существование циркуляции воздуха вокруг обтекаемого тела.

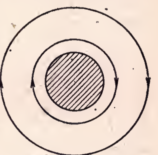


Рис. 15.4. Линии тока вязкой жидкости вокруг вращающегося цилиндра.

Чтобы лучше понять причину возникновения подъемной силы, рассмотрим вначале обтекание вращающегося цилиндра равномерным потоком воздуха. Если бы цилиндр не вращался, то из-за малой вязкости воздуха картина обтекания набегающим потоком мало отличалась бы от изображенной на рис. 15.3. Вязкий воздух «прилипает» к поверхности цилиндра. Поэтому при вращении цилиндр увлекает прилегающие слои воздуха, вызывая его циркуляцию. Если бы набегающего потока не было, то вследствие вязкости, пренебрегать которой здесь нельзя, картина линий тока воздуха вокруг вращающегося цилиндра имела бы вид, показанный на рис. 15.4. При этом чем дальше от цилиндра, тем меньше скорость увлекаемого цилиндром воздуха. При обтекании потоком воздуха вращающегося цилиндра происходит наложение картин 15.3 и 15.4. В тех местах, где скорости поступательного движения воздуха с потоком и вращения вместе с цилиндром совпадают по направлению, результирующая скорость воздуха превосходит скорость потока, набегающего на цилиндр. Там, где обусловленная вращением скорость воздуха направлена противоположно набегающему потоку, результирующая скорость воздуха меньше скорости потока. В результате получается картина обтекания набега-

ющим воздухом вращающегося цилиндра, изображенная на рис. 15.5: скорость воздуха снизу цилиндра меньше, а давление, следовательно, больше, чем сверху. Возникает подъемная сила. Это явление называется эффектом Магнуса. Его легко наблюдать экспериментально при скатывании с наклонной плоскости легкого цилиндра из плотной бумаги (рис. 15.6). Направленная перпендикулярно скорости поступательного движения цилиндра подъемная сила приводит к резкому увеличению крутизны траектории — цилиндр, падая, заворачивает под стол.

Эффект Магнуса проявляется при полете закрученного футбольного или теннисного мяча при резаных ударах.

Итак, циркуляция воздуха вокруг твердого тела, находящегося в потоке, приводит к появлению подъемной силы. В эффекте Магнуса циркуляция возникает за счет вращения цилиндра. В других случаях циркуляция может возникнуть и без вращения тела. Так, например, циркуляция возникает при обтекании вязкой жидкостью несимметричного тела.

Отметим, что циркуляция не может возникать в идеальной жидкости, где вообще не существует касательных напряжений между различными слоями жидкости. Роль вязкости в образовании циркуляции можно проиллюстрировать следующим эффектным опытом. Пусть в потоке жидкости на дне русла имеется углубление, как показано на рис. 15.7. При отсутствии вязкости могло бы существовать такое движение, при котором жидкость в углублении была бы неподвижной (рис. 15.7, *а*). При этом скорость жидкости менялась бы скачком на параллельной дну русла поверхности NN'. В реальной жидкости при скольжении придонного слоя над неподвижной водой в яме благодаря вязкости возникает касательная сила, которая приводит верхний слой воды в яме в движение в направлении потока. Но движение воды в яме ограничено стенками, и в результате в ней образуется система вращающихся «сцепленных шестерен».

Вязкость воздуха приводит к возникновению циркуляции вокруг крыла самолета. Опыт показывает, что циркуляция вокруг крыла возникает следующим образом. Вблизи острой задней кромки крыла возникают вихри, в которых вращение воздуха происходит против часовой стрелки (рис. 15.8). Эти вихри увеличиваются, отрываются

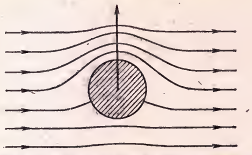


Рис. 15.5. Обтекание вращающегося цилиндра набегающим потоком.

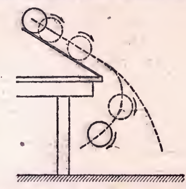


Рис. 15.6. Эффект Магнуса при скатывании легкого цилиндра с наклонной плоскости.

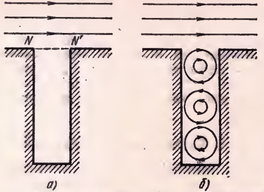


Рис. 15.7. Линии тока жидкости в русле с углублением.

от крыла и уносятся набегающим потоком воздуха. При этом остальная масса воздуха вблизи крыла начинает совершать вращение в противоположную сторону, образуя циркуляцию вокруг крыла по часовой стрелке. Циркуляционный поток, складываясь с набегающим, ускоряет движение воздуха над крылом и замедляет под крылом. В результате картина обтекания крыла потоком принимает вид, изображенный на рис. 15.9, давление воздуха над крылом понижается, под крылом повышается, что и приводит к возникновению подъемной силы.

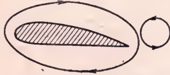


Рис. 15.8. К возникновению циркуляции воздуха вокруг крыла самолета.

При движении тела в вязкой жидкости или газе на тело, наряду с подъемной силой, действует сила лобового сопротивления. Можно указать две причины возникновения этой силы. Во-первых, вклад в лобовое сопротивление дают касательные силы внутреннего трения, действующие со стороны потока жидкости на «прилипший» к поверхности тела пограничный слой. Во-вторых, лобовое сопротивление возникает вследствие различия в силах давления на переднюю и заднюю части тела вследствие несимметричности картины обтекания вязкой жидкостью даже симметричного тела.

Рис. 15.9. Линии тока при обтекании крыла.

§ 16. Метод анализа размерностей

Заканчивая изучение механики, познакомимся еще с одним методом исследования физических процессов — так называемым методом анализа размерностей.

Рассмотрнм задачу, ответ на которую иам хорошо нзвестеи: с какой скоростью упадет на землю тело, свободио падающее без начальной скоростн с некоторой высоты h, еслн сопротивленнем воздуха можно преиебречь? Вместо того, чтобы иепосредственно определять эту скорость, пользуясь соотношеннями кннематикн, попробуем рассуждать следующнм образом. От чего вообще может зависеть величииа этой скоростн? Довольно очевидно, что от высоты h н от ускореиия свободного падення g она иепремеиио должна зависеть. Поколебавшнсь, мы можем включить в число велнчнн, от которых зависнт скорость падеиия, н массу тела m, хотя вообще-то легко сообразить, что от массы завнсимости быть не должно. Итак, предположим, что скорость падення завнснт от h, g и m:

$$v = f(h, g, m). \tag{16.1}$$

Какой вид может нметь функция f?

Ответнть на этот вопрос можио с помощью аиализа размериостей. В любой снстеме еднниц нмеется иесколько физическнх велнчнн, для которых еднинцы выбраны произвольио н счнтаются основными. В снстеме единиц СГС (а для механнческих велнчнн н в СИ) в качестве осиовиых выбраны еднницы длины L, временн T н массы M. Единицы всех остальных физнческнх величии выражаются через основиые. Напрнмер, еднннца скорости выражается через основные единицы длнны и времеии как LT^{-1}. Выраженне едииицы любой фнзнческой величииы в определеииой системе едииц через основиые едниицы этой снстемы иазывается размериостью даииой физической величины.

Поскольку складывать можно только величииы однмаковой размерностн, то после иекоторого раздумья можно для искомой фуикции f предложить такую формулу:

$$v = Ch^x g^y m^z, \tag{16.2}$$

где C — некоторое постояииое число (безразмерная постояииая), а x, y и z — неизвестные чнсла, которые следует определнть. Теперь учтем то обстоятельство, что если формула (16.2) правнльна, то размериость ее левой части должиа совпадать с размериостью правой. Размериость скоростн есть LT^{-1}, размерность высоты h есть L, размерность ускорення свободного падення g равна LT^{-2}, н,

наконец, размерность массы m равна M. Поскольку постоянная C безразмерна, то формуле (16.2) соответствует следующее равенство размерностей:

$$LT^{-1} = L^x (LT^{-2})^y M^z. \qquad (16.3)$$

Это равенство должно выполняться тождественно, независимо от того, каковы численные значения h, g и m. Поэтому следует приравнять показатели степеней при L, T и M:

$$\begin{array}{lll} L & 1 = x + y, & \\ T & -1 = -2y, & (16.4) \\ M & 0 = z. & \end{array}$$

Решая систему (16.4), находим $z = 0$, $y = 1/2$, $x = 1/2$, и поэтому формула (16.2) принимает вид

$$v = Ch^{1/2}g^{1/2}m^0 = C\sqrt{gh}. \qquad (16.5)$$

Истинная величина скорости равна, как известно, $\sqrt{2gh}$. Итак, изложенный подход дал нам возможность определить правильный вид зависимости v от h, g и m и не дал возможности найти значение постоянной C, которая на самом деле равна $\sqrt{2}$.

Теперь сразу возникает довольно много вопросов. Прежде всего, означает ли успех в разобранном примере, что это действительно универсальный способ нахождения вида зависимости между различными физическими величинами? Если да, то как определить, от каких параметров зависит интересующая нас величина? Если найдены эти параметры, то всегда ли искомая зависимость выражается формулой вида (16.2), всегда ли система уравнений типа (16.4) позволит однозначно определить показатели степеней всех величин, входящих в формулу (16.2)? Как определить значение численной постоянной C? И т. д. Не будем пытаться сразу получить исчерпывающие ответы на эти вопросы, а попытаемся постепенно выяснить все, что нас интересует, рассматривая конкретные примеры.

Первый вопрос, пожалуй, самый трудный. Судить об универсальности метода можно, только детально изучив его во всех отношениях. Поэтому мы пока оставим его в покое, разумно предположив, что, по крайней мере, всегда можно попытаться воспользоваться изложенным подходом.

Несколько яснее обстоит дело со вторым вопросом. Таблицу основных параметров, определяющих изучаемое явление, всегда можно составить, если известны описывающие его физические законы. В ряде случаев определяющие явление параметры можно указать и тогда, когда физические законы неизвестны. При определении системы параметров нужно, как и при составлении уравнений на основе физических законов, упростить, схематизировать изучаемое явление. Как правило, для использования метода анализа размерностей нужно знать меньше, чем для составления уравнений движения.

Если число параметров, определяющих изучаемое явление, больше числа основных единиц, на которых построена выбранная система единиц, то все показатели степеней, разумеется, не могут быть определены. В этом случае полезно прежде всего определить все независимые безразмерные комбинации выбранной системы параметров. Тогда искомая величина будет определяться произведением какой-либо комбинации параметров, имеющей нужную размерность, на некоторую функцию безразмерных параметров. Легко видеть, что в разобранном примере из величин h, g и m безразмерную комбинацию составить нельзя. Поэтому формула (16.2) исчерпывает все возможные случаи.

Рассмотрим теперь такую задачу: определить дальность горизонтального полета пули, выпущенной с начальной скоростью v в горизонтальном направлении на высоте h от земной поверхности. Число параметров, от которых может зависеть искомая величина, равно четырем: h, v, g и масса пули m. Поэтому полное решение задачи невозможно: система единиц СГС (а для механики и система СИ) построена только на трех основных единицах *). Найдем прежде всего все безразмерные параметры γ, которые можно сконструировать из h, v, g и m:

$$\gamma = h^x v^y g^z m^u. \qquad (16.6)$$

*) Если заранее (из каких-либо дополнительных соображений) учесть, что от массы пули искомая величина не зависит, то может показаться, что число параметров равно числу основных единиц и задача становится разрешимой. Однако это не так, ибо, исключая массу пули из числа возможных параметров, мы ограничиваем задачу рамками кинематики, а все кинематические величины содержат только две основные единицы — длину и время.

Этому выражению соответствует следующее равенство размерностей:

$$1 = L^x (LT^{-1})^y (LT^{-2})^z M^u. \qquad (16.7)$$

Отсюда получаем следующую систему уравнений:

$$\begin{array}{ll} L & 0 = x + y + z, \\ T & 0 = -y - 2z, \\ M & 0 = u. \end{array} \qquad (16.8)$$

Решая систему уравнений (16.8), найдем

$$u = 0, \quad y = -2z, \quad x = z. \qquad (16.9)$$

Теперь для искомого безразмерного параметра получаем

$$\gamma = h^z v^{-2z} g^z m^0 = \left(\frac{hg}{v^2}\right)^z. \qquad (16.10)$$

Таким образом, единственный независимый безразмерный параметр в рассматриваемой задаче — это hg/v^2. Теперь уже не нужно искать дальность полета, имеющую размерность длины, в виде (16.2): все равно, как мы видели, однозначно ее определить не удастся. Достаточно найти какой-либо параметр, имеющий размерность длины, например сам параметр h. Тогда общее выражение для дальности полета по горизонтали S можно записать в виде

$$S = hf\left(\frac{hg}{v^2}\right), \qquad (16.11)$$

где f — пока неизвестная функция безразмерного параметра $\gamma = hg/v^2$. Метод анализа размерностей в изложенном виде не позволит определить вид этой функции. Нужно привлекать какие-то дополнительные соображения. Например, из опыта нам может быть известно, что искомая дальность полета пропорциональна горизонтальной скорости пули v. Тогда функция f немедленно определяется — скорость v должна стоять в первой степени в числителе, т. е.

$$f(\gamma) = C\gamma^{-1/2}, \qquad (16.12)$$

и для S получаем

$$S = Ch\left(\frac{hg}{v^2}\right)^{-1/2} = Cv\sqrt{\frac{h}{g}}, \qquad (16.13)$$

что с точностью до постоянного множителя совпадает с правильным ответом $S = v\sqrt{2h/g}$.

Подчеркнем, что при таком способе определения вида функции f нам достаточно знания экспериментально установленной зависимости дальности полета S не от всех параметров, а только от одного из них.

Но это не единственный возможный путь определения функции f. Можно немного изменить рассуждения и определить f из соображений размерности. Действительно, если бы число основных единиц, на которых построена система единиц, равнялось бы не трем, а четырем, то равенство размерностей позволило бы полностью определить зависимость дальности полета S от h, v, g и m. Откуда же взять четвертую независимую единицу? До сих пор при записи формул размерностей не делалось различия между единицами длины в горизонтальном и вертикальном направлениях. Однако такое различие можно ввести и обозначить размерность единицы длины по горизонтали через L_x, а по вертикали — через L_y. Тогда размерность дальности полета по горизонтали S будет L_x, а размерность высоты h — L_y. Размерность горизонтальной скорости v будет L_xT^{-1}, а размерность ускорения свободного падения g — L_yT^{-2}. Теперь, глядя на формулу (16.11), мы видим, что единственный способ получить размерность L_x в числителе правой части заключается в том, чтобы считать $f(\gamma)$ пропорциональной $\gamma^{-1/2}$. Мы снова приходим к формуле (16.13). Разумеется, теперь, имея четыре основные единицы L_x, L_y, T и M, можно и непосредственно попытаться сконструировать величину нужной размерности из четырех параметров h, v, g и m:

$$S = Ch^a v^b g^c m^d. \qquad (16.14)$$

Равенство размерностей левой и правой частей

$$L_x = L_y^a (L_xT^{-1})^b (L_yT^{-2})^c M^d \qquad (16.15)$$

приводит к системе уравнений для показателей степеней a, b, c и d:

$$\begin{array}{ll} L_x & 1 = b, \\ L_y & 0 = a + c, \\ T & 0 = -b - 2c, \\ M & 0 = d. \end{array} \qquad (16.16)$$

Решая эту систему, находим

$$d = 0, \quad b = 1, \quad c = -1/2, \quad a = 1/2 \qquad (16.17)$$

и получаем формулу (16.13). Такое разложение размерности длины по взаимно перпендикулярным направлениям получило название «векторных единиц длины». Их использование, как мы видим, существенно увеличивает возможности метода анализа размерностей.

Анализ размерностей является одним из универсальных методов исследования физических явлений и очень широко используется. Великий физик Энрико Ферми часто утверждал, что действительно понимающие природу того или иного явления должны уметь получать основные соотношения из соображений размерности.

Особенно широкое применение получил метод анализа размерностей при изучении движения вязкой жидкости и газа и при изучении движения твердых тел в жидкости и газе. В этом наиболее сложном разделе динамики, где очень многие вопросы не решены по настоящее время, метод анализа размерностей зачастую оказывается единственным подходом, позволяющим теоретически осмыслить результаты различных экспериментальных исследований. Число физических величин, определяющих то или иное явление, иногда настолько велико, что оказывается практически невозможным не только решить уравнения движения, но и даже составить их. Но и в таких случаях анализ размерностей позволяет найти некоторые основные соотношения между этими величинами, что фактически означает уменьшение числа независимых параметров. Сочетание метода размерностей с экспериментом дает возможность получать частичную информацию о свойствах изучаемых систем и в очень сложных случаях, когда другими методами ее вообще получить не удается.

В качестве примера использования анализа размерностей рассмотрим одну из простейших задач этого сложного раздела механики — задачу о сопротивлении, испытываемом твердым телом при движении в вязкой среде — жидкости или газе. Сопротивление движению в вязкой среде зависит от большого числа параметров, относительная роль которых меняется в зависимости от скорости движения тела.

При небольших скоростях движения основной вклад в сопротивление определяется вязкостью жидкости, а ее плотностью и сжимаемостью можно пренебречь. При больших скоростях движения определяющую роль играет как раз плотность жидкости, а не ее вязкость. И, наконец, если скорость движения твердого тела становится сравнимой со скоростью звука в жидкости, то при расчете силы сопротивления необходимо учитывать сжимаемость среды. Таким образом, набор параметров, определяющих величину сопротивления движению, включает в себя скорость тела v, плотность среды ρ, ее вязкость η, сжимаемость $\beta = \frac{1}{V}\left|\frac{\Delta V}{\Delta p}\right|$ и линейные размеры тела l_x, l_y и l_z. Будем считать, что тело обладает симметрией относительно оси z, а его скорость v направлена вдоль оси симметрии. Тогда тело характеризуется линейным размером l вдоль оси симметрии и площадью поперечного сечения S. Будем использовать векторные единицы длины, приписывая разные размерности L_x, L_y и L_z для расстояний вдоль осей x, y и z.

Рассмотрим случай, когда скорость движения тела много меньше скорости звука в среде. При этом сжимаемостью среды можно пренебречь, и мы имеем пять параметров: v, ρ, η, l, S. Выражения для размерностей этих величин имеют следующий вид:

$$\begin{array}{ll} v & L_z T^{-1}, \\ \rho & (L_x L_y L_z)^{-1} M, \\ \eta & L_z^{-1} T^{-1} M, \\ l & L_z \\ S & L_x L_y. \end{array} \qquad (16.18)$$

Прежде всего определяем безразмерные параметры, которые можно составить из этих величин:

$$\gamma = v^x \rho^y \eta^z S^u l^w. \qquad (16.19)$$

Выражение (16.19) приводит к следующему равенству размерностей:

$$1 = (L_z T^{-1})^x (L_x^{-1} L_y^{-1} L_z^{-1} M)^y (L_z^{-1} T^{-1} M)^z (L_x L_y)^u L_z^w. \qquad (16.20)$$

Система уравнений для показателей степеней x, y, z, u и w имеет вид

$$\begin{aligned} &L_x && 0=-y+u, \\ &L_y && 0=-y+u, \\ &L_z && 0=x-y-z+w, \\ &T && 0=-x-z, \\ &M && 0=y+z. \end{aligned} \tag{16.21}$$

Решая ее, находим $x=y=u=-z$, $w=z$. Следовательно, безразмерный параметр γ, согласно (16.19), имеет вид

$$\gamma=\left(\frac{l\eta}{S\rho v}\right)^z. \tag{16.22}$$

Из выражения (16.22) видно, что в рассматриваемом случае имеется всего один независимый безразмерный параметр $l\eta/(S\rho v)$. Поэтому выражение для силы сопротивления равно произведению какой-либо комбинации из величин v, ρ, η, S и l, имеющей размерность силы, направленной вдоль оси z, на некоторую функцию безразмерного параметра $l\eta/(S\rho v)$. Как легко убедиться, произведение $v^2\rho S$ как раз и имеет размерность силы, направленной по оси z. Поэтому выражение для силы сопротивления F записывается в виде

$$F=v^2\rho S f\left(\frac{l\eta}{S\rho v}\right). \tag{16.23}$$

Эта формула позволяет сделать очень интересные выводы о силе сопротивления. Пусть скорость движения настолько мала, что определяющую роль в сопротивлении играет вязкость жидкости η. Так как сила сопротивления при этом пропорциональна вязкости, то

$$f\left(\frac{l\eta}{S\rho v}\right)\to C\frac{l\eta}{S\rho v},$$

и выражение (16.23) принимает вид

$$F=Cvl\eta, \tag{16.24}$$

где C — некоторая постоянная. Сила сопротивления пропорциональна скорости движения тела, вязкости и линейному размеру тела в направлении движения. Она оказы-

вается не зависящей от плотности жидкости и от величины поперечного сечения тела.

При большей скорости определяющей становится не вязкость жидкости, а ее плотность. Для того чтобы сила сопротивления не зависела от вязкости, нужно, чтобы функция f стремилась к постоянному значению. Формула (16.23) при этом принимает вид

$$F = C_1 v^2 \rho S, \qquad (16.25)$$

где C_1 — новая постоянная. Как и можно было ожидать из качественных соображений, сопротивление в этом случае определяется поперечным сечением тела и не зависит от размеров тела вдоль направления движения.

2. МОЛЕКУЛЯРНАЯ ФИЗИКА И ТЕРМОДИНАМИКА

ОСНОВЫ МОЛЕКУЛЯРНО-КИНЕТИЧЕСКОЙ ТЕОРИИ

§ 1. Броуновское движение. Два подхода к описанию макроскопических систем

Основное положение молекулярно-кинетической теории строения вещества, вытекающее из опытных фактов, заключается в том, что все макроскопические тела состоят из атомов и молекул, находящихся в состоянии непрерывного хаотического теплового движения. Наиболее убедительным опытным фактом, наглядно подтверждающим хаотический характер теплового движения и зависимость интенсивности этого движения от температуры, является броуновское движение.

Наблюдая в микроскоп за движением взвешенных в жидкости мелких частиц, можно обнаружить, что каждая частица совершает хаотическое движение. Представление о характере блуждания частицы можно получить, если фиксировать ее положение в поле зрения измерительного микроскопа через равные промежутки времени. Соединив последовательные положения частицы прямыми, мы получим ломаную линию, подобную изображенной на рис. 1.1. Направления соседних участков ломаной линии составляют всевозможные углы друг с другом, так что не удается подметить никакой закономерности в изменении направления ломаной линии. Чем короче промежутки времени, через которые фиксируется положение частицы, тем более изломанной будет выглядеть «траектория» частицы: точки A, B, C, . . . фиксируют положение частицы через 30 секунд, а точки 1, 2, 3, . . . , соединенные пунктиром, фиксируют ее положение через каждые 5 секунд.

Если наблюдать за движением сразу нескольких взвешенных в жидкости частиц, то можно заметить, что они

движутся то в одну сторону, то в противоположные, то под углом друг к другу. Отсюда можно сделать вывод, что наблюдаемое броуновское движение не связано с перемещением потоков жидкости, так как в этом случае соседние частицы всегда перемещались бы вместе. На опыте же никакой согласованности в движении соседних частиц не наблюдается, движутся они совершенно независимо друг от друга.

Изменяя температуру, при которой производится опыт, можно заметить, что с увеличением температуры интенсивность броуновского движения растет, с понижением температуры оно замирает.

Такой характер движения позволяет предположить, что броуновская частица движется под действием ударов, получаемых от молекул жидкости, в которой она находится. Если при этом считать, что тепловое движение молекул жидкости является хаотическим, то можно объяснить все наблюдаемые на опыте закономерности броуновского движения.

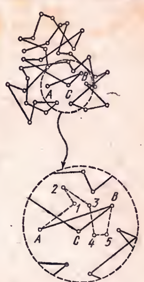


Рис. 1.1. Броуновское движение.

На первый взгляд могло бы показаться, что совершенно хаотический, беспорядочный характер ударов отдельных молекул должен был бы приводить к тому, что броуновская частица, масса которой во много раз больше массы молекулы, вообще не должна была бы заметно перемещаться. В самом деле, действие ударов, полученных броуновской частицей с одной стороны, должно полностью компенсироваться ударами с противоположной стороны. В такой ситуации, казалось бы, броуновская частица может только «дрожать» на месте. Ошибка такого рассуждения заключается в том, что случайный процесс подменяется, по существу, регулярным чередованием воздействий с противоположных сторон. Но такое чередование уже не является случайным процессом, а обладает высокой степенью упорядоченности. Степень упорядоченности такого чередо-

вания не отличается от степени упорядоченности процесса, в котором все испытываемые частицей толчки происходят в одном направлении. Если, например, результат одного толчка характеризуется некоторым расстоянием l, то результат последовательности N упорядоченных толчков пропорционален величине Nl. Если же последовательность этих N толчков носит случайный характер, то их результат пропорционален $\sqrt{N}\,l$. Покажем это.

Будем с помощью измерительного микроскопа определять расстояние, на которое броуновская частица удаляется от начала координат за время t, многократно повторяя этот опыт. Всякий раз мы будем получать разные значения этого расстояния, однако в большинстве опытов будут получаться близкие друг к другу значения, и лишь изредка будут получаться величины, заметно отличающиеся от остальных. Можно ввести среднее расстояние, на которое уходит частица от начала координат за время t, понимая под ним среднее по большому числу опытов. В однократном опыте имеется большая вероятность получить значение расстояния, близкое к среднему. Подчеркнем, что речь идет именно о расстоянии до частицы от начала координат, а направления перемещений в отдельных опытах могут быть совершенно различными — все направления равновероятны.

Задача состоит в том, чтобы найти зависимость от времени среднего расстояния, которое будем обозначать $\langle R \rangle$.

Разделим интересующее нас время наблюдения t на большое число равных малых промежутков Δt таких, что в течение каждого промежутка частица испытывает огромное число ударов со стороны молекул жидкости. По существу, такое рассуждение означает многократное повторение опыта по измерению среднего расстояния, пройденного частицей за время Δt, причем каждый раз мы совмещаем начало координат с положением частицы в конце предыдущего промежутка времени Δt. Другими словами, это такой же опыт, как и рассмотренный выше, только осуществляемый за промежуток времени Δt, а не t. Поскольку и за промежуток Δt частица испытывает огромное число ударов, все приведенные выше рассуждения остаются в силе: направление перемещения за каждый «шаг» Δt совершенно произвольно и никак не связано с направлением перемещений в другие промежутки, а расстояние, прохо-

димое частицей за Δt, будет примерно одинаковым для большинства промежутков.

Для простоты будем считать эти расстояния одинаковыми для всех шагов и обозначим через L.

Пусть в результате N таких последовательных шагов частица оказалась в точке с радиус-вектором $\boldsymbol{R}_N$. Тогда после очередного шага она попадет в точку

$$\boldsymbol{R}_{N+1} = \boldsymbol{R}_N + \boldsymbol{L}_{N+1}, \qquad (1.1)$$

где $\boldsymbol{L}_{N+1}$ — вектор перемещения за $(N+1)$-й шаг, имеющий произвольное направление и определенную величину L. Расстояние частицы от начала координат после $(N+1)$-го шага равно

$$R_{N+1} = \sqrt{R_N^2 + 2R_N L \cos\varphi + L^2}. \qquad (1.2)$$

Здесь φ — угол между векторами $\boldsymbol{R}_N$ и $\boldsymbol{L}_{N+1}$. Найти среднее значение правой части этого выражения затруднительно, ибо усреднять нужно квадратный корень, а в общем случае среднее значение функции не равно этой функции от среднего значения аргумента: $\langle f(x)\rangle \neq f(\langle x\rangle)$. Легко заметить, что если возвести (1.1) или (1.2) в квадрат:

$$R_{N+1}^2 = R_N^2 + 2R_N L \cos\varphi + L^2, \qquad (1.3)$$

то среднее значение квадрата смещения может быть легко найдено. Поэтому будем использовать для характеристики удаления броуновской частицы от начала координат не $\langle R\rangle$, а $\sqrt{\langle R^2\rangle}$. Усредняя левую и правую части (1.3) и учитывая, что угол φ с равной вероятностью принимает любые значения от 0 до 2π, т. е. $\langle \cos\varphi\rangle = 0$, получаем

$$\langle R_{N+1}^2\rangle = \langle R_N^2\rangle + L^2. \qquad (1.4)$$

Используя метод математической индукции, на основе соотношения (1.4) легко показать, что

$$\langle R_N^2\rangle = NL^2.$$

Таким образом, среднее значение квадрата смещения пропорционально числу шагов, а поскольку шаги совершаются за одинаковые промежутки времени Δt, то

$$\langle R^2\rangle = \alpha t. \qquad (1.5)$$

Это, конечно, не означает, что среднее смещение пропорционально времени. Броуновское движение частицы

таково, что средиий квадрат смещеиия растет пропорциоиальио времеии. Другими словами, квадратиый кореиь из $\langle R^2\rangle$ растет со времеием пропорциоиальио $\sqrt{t}$. Эта величииа, т. е. $\sqrt{\langle R^2\rangle}$, иазываемая средиим квадратичиым зиачеиием R, ие равиа средиему зиачеиию расстояиия $\langle R\rangle$ частицы от иачала коордииат спустя промежуток времеии t, которое мы хотели определить! Одиако можио показать, что эти величииы отличаются только постояииым миожителем. Поэтому средиее расстояиие броуиовской частицы от иачала коордииат также пропорциоиальио $\sqrt{t}$:

$$\langle R\rangle=\beta\sqrt{t}. \qquad (1.6)$$

Совершеиио очевидио, что коэффициеиты α и β в формулах (1.5) и (1.6) зависят от иитеисивиости теплового движеиия молекул жидкости, удары которых приводят к броуиовскому движеиию взвешеииой частицы, т. е. в коиечиом счете от температуры.

Изучеиие броуиовского движеиия сыграло большую роль в развитии молекулярио-кииетической теории строеиия вещества. Имеиио броуиовское движеиие ие только прииесло иеопровержимое доказательство реальиости атомов и молекул, ио и позволило впервые подсчитать количество молекул в макроскопическом объеме вещества, т. е. определить зиачеиие числа Авогадро: $N_A=6{,}02\cdot10^{23}$ моль$^{-1}$. Таким образом было окоичательио устаиовлено, что тепловая форма движеиия материи обусловлеиа хаотическим движеиием атомов или молекул, из которых состоят макроскопические тела.

Исторически при изучеиии макроскопических свойств физических систем иезависимо сложились два различиых подхода — статистический и термодииамический. Статистический подход, или статистическая механика, является молекулярио-кииетической теорией, осиоваииой иа определеииых представлеииях о строеиии вещества. Задачей статистической механики является устаиовлеиие закоиов поведения макроскопических систем, состоящих из огромиого числа частиц, иа осиове известиых дииамических закоиов поведеиия отдельиых частиц. Другими словами, статистическая мехаиика устаиавливает связь между экспериментальио измеряемыми макроскопическими величииами, характеризующими систему в целом, такими, как давлеиие,

объем, температура, напряженность электрического поля и т. д., и микроскопическими характеристиками системы, такими, как массы и заряды составляющих систему частиц, их координаты и импульсы и т. д.

Поясним сказанное на примере. Простейшей системой, состоящей из большого числа частиц, является газ, занимающий некоторый объем. С точки зрения механики состояние такой системы определяется заданием положений и скоростей всех молекул газа, число которых в макроскопическом объеме огромно. Например, всего 1 $см^3$ воздуха при нормальных условиях содержит $2{,}7\cdot10^{19}$ молекул. Из-за движения молекул механическое состояние непрерывно изменяется. Однако опыт показывает, что при неизменных внешних условиях любая макроскопическая система рано или поздно приходит в стационарное состояние, при котором, несмотря на изменение механического состояния, такие величины, как, например, температура, плотность, давление, характеризующие систему в целом, остаются неизменными. Для изолированной макроскопической системы это будет состояние теплового равновесия. Заданием таких макроскопических величин и характеризуют состояние системы в статистической механике.

Таким образом, определение состояния системы в статистической механике является гораздо менее детализированным, чем в механике, так как опирается лишь на небольшое число макроскопических параметров, измеряемых на опыте. В большинстве случаев такое сокращенное описание системы является вполне достаточным, ибо нас, как правило, совершенно не интересует детальная информация о движении отдельных молекул.

Но значения макроскопических параметров, разумеется, зависят от движения молекул, и задача статистической механики — выразить свойства системы в целом через характеристики отдельных молекул, т. е. перекинуть мост между макро- и микроскопическими описаниями системы. При этом требуется установить связь макроскопических параметров системы со средними значениями микроскопических величин и дать способ вычисления этих средних значений на основе законов движения отдельных молекул.

Так, например, для одного моля идеального газа молекулярно-кинетическая теория устанавливает связь между произведением двух макроскопических параметров газа —

давления p и молярного объема V_μ — и средним значением микроскопического параметра $\langle E \rangle$ — средней кинетической энергией хаотического движения одной молекулы E:

$$pV_\mu = \frac{2}{3} \langle E \rangle N_A. \quad (1.7)$$

В отличие от молекулярно-кинетической теории, термодинамический подход при изучении свойств макроскопических систем не опирается ни на какие модельные представления об атомно-молекулярной структуре вещества. Термодинамика представляет собой феноменологическую теорию, основанную на небольшом числе твердо установленных на опыте законов, таких, как, например, закон сохранения энергии. Основные понятия термодинамики вводятся на основе физического эксперимента, и поэтому она оперирует только макроскопическими величинами: давлением, температурой, объемом и т. п. При этом связь между различными макроскопическими параметрами конкретных систем в термодинамике устанавливается опытным путем. Например, для 1 моля газа в условиях, когда по своим свойствам он близок к идеальному, на опыте установлена следующая связь между тремя макроскопическими параметрами: давлением, молярным объемом и абсолютной температурой:

$$pV_\mu = RT. \quad (1.8)$$

Термодинамический подход отличается большой общностью и простотой. Он дает возможность решать многие конкретные задачи, не требуя никаких сведений о свойствах атомов или молекул.

Недостатком термодинамического метода можно считать то, что при его использовании остается невыявленной связь между наблюдаемым явлением и обуславливающим это явление поведением молекул. Если мы, например, термодинамическим методом установим, что металлический стержень при нагревании должен удлиняться, а растянутый резиновый жгут — сокращаться, то мы не сможем объяснить, какие особенности строения вещества приводят к такому различию в поведении при нагревании. Если же это нас не удовлетворяет и мы хотим понять, почему так происходит, то мы должны обратиться к статистической механике, так как в рамках термодинамики невозможно

вскрыть глубокий физический смысл макроскопических параметров и их связь с микроскопическими параметрами. Однако именно благодаря этому обстоятельству основные законы термодинамики, установленные на опыте, применимы ко всем макроскопическим системам, независимо от особенностей их внутренней структуры.

Статистическая механика и термодинамика долгое время развивались независимо, ибо термодинамика основывалась на экспериментальных фактах, в то время как в основе статистической механики лежали гипотезы об атомно-молекулярном строении вещества и кинетической природе теплоты, достоверность которых вызывала сомнение до тех пор, пока эти гипотезы не были подтверждены экспериментально. С тех пор отпала необходимость в резком разграничении между термодинамикой и молекулярно-кинетической теорией, и в настоящее время они фактически слились в единую науку — статистическую термодинамику.

Действительно, наиболее полные представления о свойствах систем большого числа частиц дает совместное использование термодинамики и статистической механики. Например, сравнение формул (1.7) и (1.8) дает возможность установить физический смысл макроскопического параметра — термодинамической температуры T, связав ее со средней кинетической энергией хаотического движения молекул:

$$\langle E \rangle = \frac{3}{2} kT,$$

где $k = R/N_A$ — постоянная Больцмана.

§ 2. Молекулярно-кинетическая теория идеального газа

Статистическая механика дает возможность установить связь между макроскопическими параметрами большой системы и средними значениями микроскопических величин, характеризующих отдельные молекулы. Проиллюстрируем это на простом, но важном примере уравнения состояния идеального газа.

Будем рассматривать идеальный газ как совокупность огромного числа одинаковых молекул, размеры которых пренебрежимо малы. Будем считать, что молекулы движутся по законам классической механики и взаимодействуют

между собой только во время столкновений, которые носят характер упругого удара. Газ заключен в сосуд, и в состоянии теплового равновесия никаких макроскопических движений в нем не происходит.

Рассчитаем давление, оказываемое идеальным газом на стенку сосуда. Суммарное действие молекул на поверхность можно заменить одной непрерывно действующей силой, так как молекул очень много и их столкновения со стенкой происходят очень часто. Поэтому, согласно законам динамики Ньютона, давление газа равно величине изменения перпендикулярной к стенке составляющей импульса всех

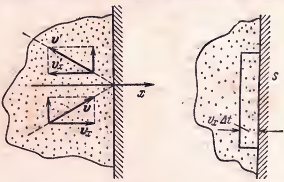


Рис. 2.1. Упругое столкновение молекулы газа со стенкой сосуда.

Рис. 2.2. К вычислению давления газа.

молекул, которые испытывают столкновения с единицей площади поверхности стенки за единицу времени. Хотя место и время удара каждой молекулы о стенку совершенно несущественно, начинать приходится с рассмотрения удара отдельной молекулы. Пока для простоты предположим, что молекулы сталкиваются со стенкой абсолютно упруго. Когда молекула отскакивает от стенки, проекция ее скорости на направление нормали к стенке меняет знак. Направим ось x по нормали к стенке (рис. 2.1). Обозначим через v_x проекцию скорости молекулы до удара, а через $v'_x = -v_x$ после удара. Изменение импульса молекулы при столкновении со стенкой $m(v'_x - v_x)$ равно $-2mv_x$, а передаваемый стенке импульс равен $2mv_x$. Это относится к единичному столкновению. Теперь рассмотрим те молекулы, у которых проекция скорости на ось x лежит в малом

интервале от v_x до $v_x+\Delta v_x$. Пусть число таких молекул в единице объема равно $\Delta n(v_x)$. За время Δt до стенки долетят и столкнутся с ней только те из них, которые находятся внутри слоя толщиной $v_x\Delta t$, прилегающего к участку стенки площади S (рис. 2.2). Промежуток времени Δt можно выбрать настолько малым, чтобы толщина слоя $v_x\Delta t$ была много меньше средней длины свободного пробега молекул. Тогда столкновений молекул между собой в этом слое практически не будет. Итак, число ударов, наносимых рассматриваемыми молекулами за время Δt, равно $\Delta n(v_x)v_xS\,\Delta t$, а передаваемый при этом стенке импульс $\Delta F\,\Delta t$ равен

$$\Delta F\,\Delta t = 2mv_x\,\Delta n(v_x)\,v_xS\,\Delta t.$$

Отсюда давление на стенку Δp, создаваемое этой группой молекул, равно

$$\Delta p = \frac{\Delta F}{S} = 2mv_x^2\,\Delta n(v_x). \qquad (2.1)$$

Полное давление, создаваемое всеми молекулами, получим, просуммировав (2.1) по всем группам молекул, скорости которых направлены к стенке, т. е. по всевозможным значениям $v_x>0$:

$$p = 2m\sum v_x^2\,\Delta n(v_x) \qquad (v_x>0). \qquad (2.2)$$

Вследствие хаотичности теплового движения в состоянии равновесия число молекул со скоростью, лежащей в интервале от v_x до $v_x+\Delta v_x$, летящих к стенке, в среднем равно числу молекул со скоростью от $-v_x$ до $-(v_x+\Delta v_x)$, летящих от стенки, т. е. $\Delta n(v_x)=\Delta n(-v_x)$. Так как под знаком суммы в (2.2) стоит квадрат проекции скорости, то сумма только по положительным значениям v_x равна половине суммы по всевозможным v_x:

$$p = m\sum v_x^2\,\Delta n(v_x). \qquad (2.3)$$

Легко сообразить, что сумма в (2.3) связана со средним значением квадрата проекции скорости молекулы на ось x. В самом деле, среднее значение v_x^2 по совокупности из n молекул определяется формулой

$$\langle v_x^2\rangle = \frac{1}{n}\sum_i v_{xi}^2 = \frac{1}{n}\left(v_{x1}^2 + v_{x2}^2 + \ldots + v_{xn}^2\right). \qquad (2.4)$$

Но в (2.3) фактически стоит та же самая сумма, только суммирование производится не по отдельным молекулам, а по группам молекул в единице объема, имеющих почти одинаковые значения v_x. Поэтому (2.3) можно переписать в виде

$$p = mn\langle v_x^2\rangle, \tag{2.5}$$

где n есть полное число молекул в единице объема или концентрация. В силу равноправия всех направлений при хаотическом тепловом движении $\langle v_x^2\rangle = \frac{1}{3}\langle v^2\rangle$, и, следовательно, формула (2.5) принимает вид

$$p = \frac{1}{3}mn\langle v^2\rangle. \tag{2.6}$$

Это и есть основное уравнение кинетической теории идеального газа. Макроскопический параметр, характеризующий газ в целом,— давление p — выражен здесь через среднее значение микроскопического параметра — квадрата скорости отдельной молекулы.

Уравнение (2.6) можно переписать в несколько ином виде, если ввести среднюю кинетическую энергию хаотического движения молекулы газа $\langle E\rangle = \frac{m}{2}\langle v^2\rangle$:

$$p = \frac{2}{3}n\langle E\rangle. \tag{2.7}$$

Подумаем теперь о тех упрощениях, которые были сделаны при выводе уравнения (2.6).

Во-первых, может сложиться впечатление, что приведенный вывод уравнения (2.6) не учитывает столкновений молекул между собой. В самом деле, при подсчете числа ударов о площадку мы считали, что все молекулы из выделенной группы со скоростями от v_x до $v_x+\Delta v_x$, лежащие в слое толщиной $v_x\Delta t$, достигают стенки в течение промежутка времени Δt. Но на самом деле некоторые из них не долетят до стенки, так как в результате столкновений могут изменить направление скорости. Но кроме таких столкновений, в результате которых молекулы «уходят» из выделенной группы, в газе происходят столкновения, пополняющие эту группу молекул. В состоянии теплового равновесия, когда в газе отсутствуют макроскопические потоки, среднее число молекул в каждой группе не ме-

няется со временем, несмотря на частые столкновения молекул. Поэтому столкновения не могут изменить среднего результата ударов молекул о стенки сосуда, выражаемого (2.6). Второе из сделанных при выводе (2.6) предположений касалось характера взаимодействия молекулы со стенкой. Предполагалось, что удар о стенку абсолютно упругий. На самом деле это предположение также является несущественным. Чтобы убедиться в этом, рассмотрим газ в прямоугольном сосуде (рис. 2.3), в котором стенки A и B обладают разными свойствами: удар молекул газа о стенку A абсолютно упругий, а о стенку B — неупругий. Если бы давление газа на эти стенки было различным, то в отсутствие трения о подставку сосуд с газом пришел бы в движение под действием внутренних сил. Так как сосуд в движение не приходит, мы должны заключить, что в состоянии теплового равновесия давление газа не зависит от характера взаимодействия его молекул со стенками. Объясняется это тем, что в стационарном состоянии молекулы не накапливаются на стенках: сколько молекул прилипает к стенке при неупругом ударе, столько же от нее и улетает, причем в среднем с такой же скоростью, так как температура стенки и газа одинакова.

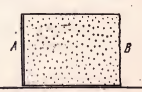


Рис. 2.3. Стенки сосуда A и B разные: удар молекул о стенку A абсолютно упругий, о стенку B — неупругий.

А вот предположение о том, что газ идеальный, т. е. его молекулы настолько малы, что их полный собственный объем мал по сравнению с объемом занимаемого газом сосуда, и что молекулы взаимодействуют только во время столкновений, оказывается существенным. Только для газа с такими свойствами и справедливо уравнение (2.6). Для реальных газов оно выполняется лишь приближенно.

Установленная на опыте связь между давлением, объемом и температурой газа приближенно описывается уравнением Менделеева — Клапейрона, которое для одного моля газа имеет вид

$$pV_\mu = RT. \tag{2.8}$$

Это уравнение выполняется тем точнее, чем ближе газ по

своим свойствам к идеальному. Для идеального газа оно было бы точным.

Преобразуем уравнение (2.7) так, чтобы его было удобно сравнивать с (2.8). Умножим (2.7) на объем V_μ, занимаемый одним молем газа. Так как число молекул в одном моле любого газа есть число Авогадро N_A, то $nV_\mu = N_A$ и

$$pV_\mu = \frac{2}{3} N_A \langle E \rangle. \qquad (2.9)$$

Теперь сравнение уравнений (2.8) и (2.9) позволяет выразить среднюю кинетическую энергию хаотического теплового движения молекул $\langle E \rangle$ через температуру газа T:

$$\langle E \rangle = \frac{3}{2} kT, \quad k = \frac{R}{N_A}. \qquad (2.10)$$

Отметим, что под $\langle E \rangle$ понимается средняя кинетическая энергия поступательного движения молекул, ибо давление газа определяется передаваемым стенке импульсом, который зависит только от скорости поступательного движения молекул. Поэтому формула (2.10) для одноатомного газа дает среднюю величину полной энергии хаотического движения, а для многоатомных газов, где тепловое движение включает в себя также вращение молекул и колебания атомов в молекулах, она дает среднее значение только кинетической энергии поступательного движения.

Самое интересное в формуле (2.10) то, что средняя энергия теплового движения молекул зависит только от температуры газа. При данной температуре средняя кинетическая энергия поступательного хаотического движения молекул не зависит ни от химического состава газа, ни от массы молекул, ни от давления газа, ни от объема, занимаемого газом.

Подставляя $\langle E \rangle$ из (2.10) в формулу (2.7), получаем следующее выражение для давления идеального газа:

$$p = nkT, \qquad (2.11)$$

где n — концентрация молекул.

В идеальном газе, где молекулы не взаимодействуют между собой, энергия всего газа есть просто сумма энергий отдельных молекул. Для одного моля одноатомного газа эта энергия U равна

$$U = \frac{3}{2} kN_A T = \frac{3}{2} RT. \qquad (2.12)$$

Эта величина не связана с движением газа как целого и называется внутренней энергией газа.

Для неидеального газа внутренняя энергия представляет собой сумму энергий отдельных молекул и энергии их взаимодействия.

Поскольку в состоянии теплового равновесия средняя кинетическая энергия молекул зависит только от температуры, то в смеси газов средние кинетические энергии молекул разных сортов одинаковы:

$$\frac{m_1 \langle v_1^2 \rangle}{2} = \frac{m_2 \langle v_2^2 \rangle}{2} = \frac{3}{2} kT. \qquad (2.13)$$

Поэтому легкие молекулы движутся в среднем быстрее тяжелых.

Полное давление смеси идеальных газов равно сумме давлений, которые имел бы каждый из газов, составляющих смесь, если удалить из сосуда остальные газы. В этом можно убедиться, буквально повторяя приведенный вывод уравнения (2.6) и учитывая, что импульс, передаваемый стенке молекулами каждого сорта, обуславливает то давление, которое создавал бы один этот газ. Этот закон был открыт опытным путем Дальтоном и носит его имя.

§ 3. Статистические распределения

Полный беспорядок, которым характеризуется тепловое движение молекул, все же имеет свои законы. Несмотря на то, что каждая молекула газа при столкновениях с другими молекулами и со стенками сосуда все время изменяет свою скорость, макроскопическое состояние газа в равновесии не меняется, что позволяет считать, что в газе существует некоторое в среднем неизменное во времени распределение молекул по скоростям. Действительно, как мы видели, при данной температуре среднее значение квадрата скорости молекул имеет определенное значение. Однако среди молекул в данный момент времени есть и быстрые, и медленные, и можно поставить вопрос: сколько в среднем в газе молекул имеет то или иное значение скорости?

Такое распределение молекул газа по скоростям устанавливается всегда, когда газ приходит в равновесие, независимо от того, каково было начальное состояние

системы. Если даже в откачанный сосуд впустить струю газа, в которой все молекулы имеют почти одинаковые по величине и направлению скорости, то спустя некоторое время в результате столкновений молекул направленное движение в газе перейдет в хаотическое, при котором все направления скоростей будут встречаться одинаково часто, а в распределении молекул по величине скорости будет наблюдаться определенная закономерность.

Равновесное состояние газа характеризуется не только распределением молекул по скоростям, но и по координатам. В отсутствие внешних полей это распределение будет однородным, т. е. газ равномерно распределяется по всему объему сосуда: в любых равных макроскопических объемах внутри сосуда в среднем находится одинаковое число молекул. А как обстоит дело при наличии действующего на молекулы поля, например поля тяжести? Хорошо известно, что давление воздуха убывает с высотой. Следовательно, убывает и концентрация молекул воздуха. Например, на высоте Эльбруса (5600 м) давление составляет лишь половину давления на уровне моря, т. е. концентрация молекул там уже вдвое меньше. Отсюда, конечно, не следует делать вывод, что на вдвое большей высоте совсем нет молекул воздуха,— самолеты летают и гораздо выше.

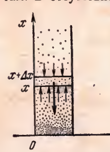


Рис. 3.1. Равновесие мысленно выделенного слоя газа в поле тяжести.

Найти закон распределения молекул газа в однородном поле тяжести с высотой можно из условия механического равновесия. Рассмотрим вертикальный столб газа с площадью основания S (рис. 3.1) и выделим в нем мысленно на высоте x слой толщиной Δx, настолько малой, чтобы плотность газа ρ можно было считать в пределах этого слоя постоянной, но в то же время эта толщина должна быть такой, чтобы внутри выделенного слоя было много молекул и можно было бы говорить о производимом ими давлении.

К этому выделенному слою газа мы применим условие механического равновесия подобно тому, как это делалось в гидростатике для слоя жидкости, где мы, используя

понятие давления, совершенно не интересовались его молекулярно-кинетической природой. Мы можем так поступать, ибо давление газа на стенку сосуда, рассматриваемое как результат передачи молекулами импульса стенке при столкновениях, и гидростатическое давление в газе на опыте измеряются одинаково, одними и теми же приборами и, следовательно, представляют собой один и тот же макроскопический параметр рассматриваемой системы.

Условие равновесия выделенного слоя газа состоит в том, что действующая на него сила тяжести уравновешивается силами давления на верхнее и нижнее основания. В проекции на ось x (рис. 3.1) это условие записывается в виде

$$p(x)S-p(x+\Delta x)S-\rho gS\,\Delta x=0.$$

Так как давление на высоте $x+\Delta x$ можно записать в виде

$$p(x+\Delta x)=p(x)+\Delta p,$$

то условие равновесия принимает вид

$$\Delta p=-\rho g\,\Delta x. \tag{3.1}$$

Входящая в (3.1) плотность газа ρ зависит от давления. Выразим ее из уравнения Менделеева — Клапейрона для произвольной массы газа M:

$$pV=\frac{M}{\mu}RT, \tag{3.2}$$

где μ — молярная масса. С помощью (3.2) получаем

$$\rho=\frac{M}{V}=\frac{\mu p}{RT}. \tag{3.3}$$

Подставляем это выражение в (3.1) и переходим к пределу при $\Delta x\to 0$. Так как предел отношения $\Delta p/\Delta x$ при $\Delta x\to 0$ есть производная dp/dx, то получаем следующее дифференциальное уравнение для функции $p(x)$:

$$\frac{dp}{dx}=-\frac{\mu g}{RT}p. \tag{3.4}$$

Это уравнение говорит о том, что производная искомой функции пропорциональна самой функции. Как известно из курса математики, единственной функцией, обладающей

таким свойством, является экспонента, и, следовательно, решение такого уравнения при постоянных g и T имеет вид

$$p(x) = C \exp\left(-\frac{\mu g x}{RT}\right). \tag{3.5}$$

Значение постоянной C определяется из условия, что давление на высоте $x=0$ равно заданной величине p_0:

$$p(x) = p_0 \exp\left(-\frac{\mu g x}{RT}\right). \tag{3.6}$$

Формулу (3.6) можно переписать в несколько ином виде, учитывая, что молярная масса μ равна произведению массы молекулы m на число Авогадро N_A:

$$p(x) = p_0 \exp\left(-\frac{m g x}{kT}\right). \tag{3.7}$$

Соотношение (3.6) или (3.7) носит название барометрической формулы. Выражаемая ею зависимость давления газа p от высоты x графически представлена на рис. 3.2. Отметим, что применимость барометрической формулы к реальной земной атмосфере весьма ограиичена, так как атмосфера практически никогда не находится в состоянии теплового равновесия и ее температура меняется с высотой. Учитывая связь между давлением газа и концентрацией молекул (2.11), из (3.7) получаем распределение молекул по высоте во внешнем поле:

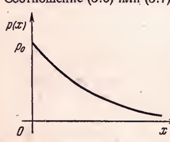


Рис. 3.2. Зависимость давления газа от высоты.

$$n(x) = n_0 \exp\left(-\frac{m g x}{kT}\right). \tag{3.8}$$

Легко заметить, что в числителе показателя экспоненты в (3.8) стоит потенциальная энергия молекулы, находящейся в поле тяжести на высоте x: $E_п(x) = mgx$, т. е. эту формулу можно переписать в виде

$$n(x) = n_0 \exp\left(-\frac{E_п(x)}{kT}\right). \tag{3.9}$$

Этот полученный на конкретном примере результат имеет весьма общий характер. Формула (3.9) дает равновесное распределение молекул в пространстве в любом потенциальном поле и носит название распределения Больцмана.

Общая теория равновесных статистических распределений была создана Гиббсом. Он показал, что в состоянии теплового равновесия закон распределения молекул по любой характеризующей их состояние величине (координате, скорости, энергии) имеет экспоненциальный характер, причем в показателе экспоненты, как и в (3.9), стоит взятое со знаком минус отношение характерной энергии молекулы к величине kT, которая пропорциональна средней кинетической энергии хаотического движения молекул. В частности, для распределения молекул газа по проекции скорости на какое-либо направление в показателе экспоненты стоит отношение зависящей от этой проекции части энергии молекулы к kT:

$$f(v_x) = a \exp\left(-\frac{E(v_x)}{kT}\right) = a \exp\left(-\frac{mv_x^2}{2kT}\right). \quad (3.10)$$

Величина $f(v_x)$ называется функцией распределения молекул по проекции скорости на ось x. Произведение $f(v_x)\Delta v_x$ равно среднему числу молекул газа в единице объема, у которых проекция скорости на ось x лежит в интервале от v_x до $v_x+\Delta v_x$, подобно тому как произведение $n(x)$ из формулы (3.9) на Δx дает среднее число молекул, x-координата которых лежит между x и $x+\Delta x$.

Остановимся подробнее на смысле функции распределения $f(v_x)$. Прежде всего подчеркнем, что бессмысленно задавать вопрос о том, сколько молекул имеют строго определенное значение v_x, например ровно 500 м/с. Скорее всего, в данный момент во всем сосуде с газом не окажется ни одной молекулы с таким значением v_x, так как число молекул газа хоть и очень велико, но все же конечно, в то время как допустимых значений v_x бесконечно много. Поэтому имеет смысл говорить только о среднем числе молекул в единице объема, значение проекции скорости которых на ось x лежит в интервале от v_x до $v_x+\Delta v_x$, например от 500 до 501 м/с.

Наглядное представление о законе распределения молекул по проекции скорости дает график функции $f(v_x)$ (3.10), который приведен на рис. 3.3. Площадь заштрихо-

ванной полоски на этом рисунке, равная $f(v_x)\Delta v_x$, дает, как мы видели, среднее число молекул газа в единице объема, у которых проекция скорости v_x лежит в указанном интервале Δv_x. Теперь легко сообразить, что полная площадь, ограниченная графиком функции $f(v_x)$ и осью v_x,

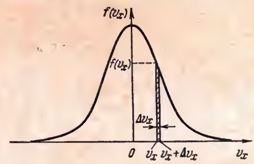


Рис. 3.3. График функции распределения $f(v_x)$.

дает число молекул, у которых v_x имеет любые значения от $-\infty$ до ∞, т. е. дает полное число молекул в единице объема. Именно из этого условия, называемого условием нормировки функции распределения, и определяется постоянная a в формуле (3.10). Для ее нахождения нужно проинтегрировать функцию $f(v_x)$ по v_x от $-\infty$ до ∞ и приравнять результат концентрации молекул n. Это дает следующее значение постоянной a:

$$a = n\sqrt{\frac{m}{2\pi kT}}. \qquad (3.11)$$

Функции распределения можно дать и несколько иную интерпретацию. Вместо того, чтобы говорить о том, сколько в среднем молекул в единице объема имеют значение v_x в заданном интервале, можно говорить о том, какова вероятность того, что наугад выбранная молекула имеет значение v_x в этом интервале. Очевидно, что эта вероятность равна отношению среднего числа таких молекул к полному числу молекул в единице объема. Обозначая ее через $q(v_x)\Delta v_x$, можем написать

$$q(v_x) = \frac{1}{n} f(v_x) = \sqrt{\frac{m}{2\pi kT}} \exp\left(-\frac{mv_x^2}{2kT}\right). \qquad (3.12)$$

Из-за полной хаотичности теплового движения в состоянии равновесия можно считать, что все направления равноправны. В противном случае в газе существовало бы некоторое преимущественное направление движения молекул и, следовательно, существовал бы направленный поток газа, чего нет в состоянии теплового равновесия. Поэтому, если выбрать любые три взаимно перпендикулярных направления x, y, z, то функции распределения молекул по проекциям скорости на эти направления будут иметь один и тот же вид (3.12).

Можно поставить вопрос, какова вероятность того, что наугад выбранная молекула газа будет иметь значения трех проекций ее скорости v_x, v_y и v_z в заданных интервалах: проекция на ось x в интервале от v_x до $v_x+\Delta v_x$, на ось y — от v_y до $v_y+\Delta v_y$, на ось z — от v_z до $v_z+\Delta v_z$. Равноправие всех направлений позволяет считать, что вероятность иметь, например, v_x в заданном интервале не зависит от того, каковы значения двух других проекций скорости молекулы. Это значит, что вероятность того, что все три проекции скорости лежат в заданных интервалах, равна произведению вероятностей для каждой из проекций в отдельности:

$$q(v_x,\ v_y,\ v_z)\,\Delta v_x\,\Delta v_y\,\Delta v_z = q(v_x)\,\Delta v_x\cdot q(v_y)\,\Delta v_y\cdot q(v_z)\,\Delta v_z.$$

Теперь с помощью (3.12) можем написать

$$q(v_x,\ v_y,\ v_z) = \left(\frac{m}{2\pi kT}\right)^{3/2}\exp\left[-\frac{m}{2kT}\left(v_x^2+v_y^2+v_z^2\right)\right]. \qquad (3.13)$$

В показателе экспоненты в (3.13) фактически стоит квадрат скорости молекулы, и показатель равен отношению ее кинетической энергии к kT. Таким образом, эта функция распределения зависит только от величины скорости, но не от ее направления. Распределение по скоростям, как и следовало ожидать, оказывается равномерным по всем направлениям, т. е. изотропным в пространстве.

Теперь легко получить выражение для вероятности того, что наугад выбранная молекула имеет величину скорости в заданном интервале от v до $v+\Delta v$, независимо от того, как эта скорость направлена. Для этого нужно найти, чему соответствует произведение $\Delta v_x \Delta v_y \Delta v_z$ для всех молекул, величина скорости которых лежит в заданном интервале от v до $v+\Delta v$. Если построить систему координат v_x, v_y, v_z (рис. 3.4), то легко видеть, что молекулам

с одинаковым значением величины скорости v соответствуют точки на поверхности сферы радиуса v с центром в начале координат. Молекулам со скоростями в интервале от v до $v+\Delta v$ соответствует шаровой слой толщины Δv. Поэтому произведению $\Delta v_x \Delta v_y \Delta v_z$ в рассматриваемом случае соответствует объем этого шарового слоя, равный произведению площади поверхности сферы на толщину слоя, т. е. $4\pi v^2 \Delta v$. Выражение для вероятности того, что молекула имеет величину скорости в заданном интервале, равно произведению (3.13) на $4\pi v^2 \Delta v$:

$$q(v)\,\Delta v = 4\pi \left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left(-\frac{mv^2}{2kT}\right) v^2\,\Delta v. \qquad (3.14)$$

Если нас интересует среднее число молекул в единице объема с такими значениями величины скорости, то мы должны умножить (3.14) на концентрацию газа n:

$$f(v)\,\Delta v = 4\pi n \left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left(-\frac{mv^2}{2kT}\right) v^2\,\Delta v. \qquad (3.15)$$

Зависимость функции $f(v)$ от скорости показана на рис. 3.5. Эта функция имеет максимум при значении $v=\sqrt{2kT/m}$, называемом наиболее вероятной скоростью. Площадь, ограниченная графиком функции и осью v, дает полное число молекул в единице объема n.

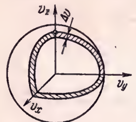


Рис. 3.4. Шаровой слой в пространстве скоростей.

При повышении температуры кривая распределения молекул по скоростям деформируется так, как показано на рис. 3.6. Максимум кривой смещается при увеличении T в область бóльших значений v. Максимальное значение $f(v)$ при этом убывает, так что площадь под кривой остается неизменной. Приведенные выше функции распределения молекул газа по скоростям были впервые получены Максвеллом и носят его имя.

Экспериментальное измерение скоростей молекул и проверка закона распределения Максвелла осуществляются различными методами, использующими молекулярные пучки. Один из первых таких опытов был проделан Штерном в 1920 году.

Знание статистических функций распределения дает возможность вычислять средние значения микроскопических параметров. Например, с помощью функции распределения Больцмана можно найти положение центра масс

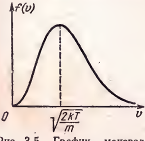


Рис. 3.5. График максвелловского распределения молекул газа по скоростям.

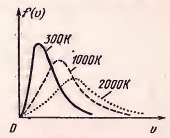


Рис. 3.6. Функции распределения молекул по скоростям при разных температурах.

газа в поле тяжести. С помощью функции распределения Максвелла вычисляются средние значения величин, зависящих от скорости молекулы. Если вычислить среднее значение квадрата скорости $\langle v^2 \rangle$, то мы получим уже приведенную ранее величину $3kT/m$. Для среднего значения абсолютной величины скорости расчет с максвелловской функцией распределения дает $\langle v \rangle = \sqrt{8kT/\pi m}$.

§ 4. Флуктуации

Тепловое равновесие — это всегда динамическое равновесие. Тепловое движение атомов или молекул, из которых состоит макроскопическая система, никогда не прекращается. Поэтому макроскопические величины, характеризующие систему в целом, строго говоря, никогда не остаются постоянными, а испытывают малые беспорядочные колебания вблизи некоторых средних значений. Такие хаотические отклонения от средних значений тех или иных величин, происходящие в течение малых промежутков времени, называются флуктуациями. Относительная величина флуктуаций тем больше, чем меньше размеры изучаемой системы.

Яркий пример флуктуаций — это дрожание зеркальца чувствительного гальванометра. Макроскопическая система,

состоящая из подвижной катушки гальванометра, подвешенной на упругой кварцевой нити, в состоянии механического равновесия была бы совершенно неподвижной, если бы не тепловое движение. Удары молекул воздуха, совершающих тепловое движение, приводят к тому, что угол поворота зеркальца испытывает хаотические колебания вблизи положения механического равновесия. Фактически это то же броуновское движение, которое отличается от рассмотренного в § 1 движения взвешенной в жидкости частицы только тем, что здесь рассматривается не поступательное, а вращательное движение вблизи устойчивого, а не безразличного положения равновесия. Интенсивность такого движения зависит от температуры, оно принципиально неустранимо и ставит предел чувствительности измерительной аппаратуры.

Основные закономерности флуктуаций можно подметить, рассматривая простейший пример пространственного распределения молекул идеального газа внутри сосуда в состоянии теплового равновесия. В среднем газ равномерно заполняет весь сосуд, т. е. концентрация молекул всюду одинакова. Разделим мысленно сосуд на две равные части. Пусть число молекул слева равно n_1, справа — n_2. Сумма n_1+n_2 есть полное число молекул в сосуде. В равновесии в среднем $n_1=n_2$, но так как это равновесие динамическое, то в каждый момент времени вследствие хаотического движения это верно лишь приближенно, потому что молекулы непрерывно переходят из одной половины сосуда в другую. В принципе ничто не мешает им вообще в какой-то момент собраться в одной половине сосуда. Однако такое событие будет крайне маловероятным. Чтобы убедиться в этом, рассмотрим, какими способами молекулы могут распределяться между половинами сосуда. Одну молекулу можно разместить в сосуде двумя способами — либо в левой, либо в правой его половине. Очевидно, что вероятность найти молекулу в какой-то одной половине сосуда равна 1/2, если, конечно, сосуд разделен на строго равные части. Две молекулы можно распределить в сосуде четырьмя (2^2) способами, ибо для каждого из двух способов распределения одной молекулы существует два способа распределения другой. Для трех молекул число способов распределения равно восьми (2^3), так как для каждого из четырех возможных вариантов распределения первых

двух молекул есть две возможности распределения третьей молекулы, и т. д. В случае N молекул число способов распределения равно 2^N.

Будем считать, что вероятность нахождения любой наугад взятой молекулы в определенной половине сосуда не зависит от того, где в этот момент находятся все остальные молекулы. Это верно для идеального газа, молекулы которого не взаимодействуют между собой и имеют пренебрежимо малые размеры, так что их собственный объем значительно меньше объема сосуда. В этом случае каждый из 2^N способов размещения N молекул в сосуде имеет одну и ту же вероятность, равную $(1/2)^N$. Разъясним это подробнее. Молекулы газа находятся в тепловом движении, и их расположение в сосуде непрерывно изменяется. Предположим, что мы можем делать мгновенные «фотографии» положения молекул в сосуде. На каждой такой «фотографии» мы увидим один из мыслимых способов размещения молекул. Размещения на любых двух снимках считаются одинаковыми, если в какой-либо половине сосуда, например левой, на обоих снимках находятся одни и те же молекулы. Утверждение о равной вероятности всех 2^N размещений означает, что какое-либо размещение будет встречаться на снимках не чаще и не реже других, в среднем один раз в каждой серии, содержащей 2^N «фотографий».

Разместить все N молекул в одной половине сосуда можно только одним способом. Поэтому вероятность того, что весь газ соберется в одной половине сосуда, равна $(1/2)^N$. При больших N это очень малая величина. Например, для газа, содержащего всего 100 молекул, вероятность такого события равна $2^{-100} \approx 10^{-30}$. В среднем только на одном из 10^{30} снимков мы бы увидели, что одна из половин сосуда пуста. Чтобы представить себе, насколько велико это число, вспомним, что, по преданию, изобретатель шахматной игры потребовал в награду «всего» $2^{64} \approx 10^{19}$ зерен пшеницы. Это оказалось больше, чем количество зерен, которое может уместиться в амбаре высотой 5 м и шириной 20 м, протянувшемся от Земли до Солнца. И это для газа, состоящего всего из ста молекул! А для одного моля газа, где число молекул составляет $6 \cdot 10^{23}$, вероятность всем молекулам оказаться в одной половине сосуда настолько мала, что просто невозможно найти для нее подходящее сравнение в мире образов, доступных человеческому воображению.

Подсчитаем теперь вероятность того, что в определенной половине сосуда находится заданное число n молекул. Эта вероятность равна отношению числа способов размещения, при которых в выбранной половине сосуда находятся любые n молекул, к 2^N— полному числу возможных размещений.

Число таких размещений, очевидно, равно числу способов, которыми можно выбрать n молекул из совокупности N молекул, т. е. числу сочетаний из N по n. Это число сочетаний равно

$$C_N^n = \frac{N!}{n!\,(N-n)!}. \qquad (4.1)$$

Поэтому вероятность P_n того, что в какой-нибудь, например левой, половине сосуда находится n молекул, равна

$$P_n = \frac{N!}{2^N n!(N-n)!}. \qquad (4.2)$$

Теперь с помощью (4.2) легко рассчитать, как часто на наших «фотографиях» будут встречаться те или иные распределения молекул. Для газа из ста молекул вероятность того, что молекулы распределятся между половинами сосуда строго поровну, приблизительно равна 1/18, т. е. такое распределение встречается в среднем 1 раз на каждые 18 снимков.

Итак, строго равномерное распределение молекул по объему встречается не так уж и часто. Но если мы подсчитаем вероятности таких распределений молекул, при которых в одной половине сосуда будет на две, четыре и вообще на небольшое число молекул больше, чем в другой, и просуммируем эти вероятности, то мы увидим, что вероятность появления любого из таких распределений будет близка к единице, а очень неравномерные распределения, при которых число молекул справа и слева различается сильно, будут встречаться крайне редко. Представление о характере распределения вероятностей (4.2) дает рис. 4.1, построенный для газа из $N=20$ молекул. По оси абсцисс отложено число молекул в одной (например, левой) половине сосуда, а по оси ординат — вероятности соответствующих размещений молекул. Приведенное распределение вероятностей можно охарактеризовать плавной кривой — огибающей, которая показана пунктиром на рис. 4.1. С увеличением полного

числа молекул газа эта кривая становится все более и более острой. Для $N=100$ она приведена на рис. 4.2.

Теперь мы можем сделать следующий общий вывод: если полное число молекул в изучаемой системе велико,

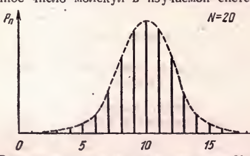


Рис. 4.1. Вероятность того, что в сосуде, содержащем 20 молекул газа, n из них находятся в одной половине сосуда, а остальные $20-n$ в другой.

то флуктуации плотности, соответствующие существенно неоднородному распределению, почти никогда не возникают. Описанные выше закономерности справедливы не только для флуктуаций плотности, но и для флуктуаций других макроскопических параметров. Если число частиц в системе очень велико, то относительная величина флуктуаций любого параметра очень мала по сравнению со средним значением этого параметра, и ею почти всегда можно пренебречь. Поэтому мы обычно не осознаем факта существования флуктуаций, когда имеем дело с большими макроскопическими системами. Но если система достаточно мала, то флуктуации могут быть легко обнаружены и часто имеют большое значение.

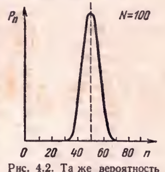


Рис. 4.2. Та же вероятность для сосуда, содержащего 100 молекул газа.

Как уже упоминалось, тепловые флуктуации ставят предел чувствительности любых измерительных приборов. При измерении малых значений физических величин или слабых сигналов флуктуации в чувствительном элементе изме-

рительного прибора могут оказаться сравнимыми с измеряемой величиной. Флуктуации, возникающие в самом измерительном приборе, затрудняют измерения и поэтому называются шумами.

ОСНОВЫ ТЕРМОДИНАМИКИ

§ 5. Первый закон термодинамики. Тепловые двигатели

Любая изолированная макроскопическая система с течением времени приходит в состояние термодинамического равновесия, в котором, если отвлечься от флуктуаций, характеризующие ее макроскопические параметры не меняются. Процесс перехода системы из неравновесного состояния в состояние равновесия называется релаксацией. Длительность этого процесса характеризуется временем релаксации.

Мы уже отмечали, что тепловое равновесие с точки зрения молекулярно-кинетической теории представляет собой динамическое равновесие: механическое состояние системы вследствие теплового движения молекул непрерывно изменяется, но термодинамическое состояние не меняется. Динамический характер теплового равновесия отчетливо виден на примере равновесия жидкости и ее насыщенного пара: в такой системе непрерывно идет процесс испарения жидкости и обратный ему процесс конденсации пара. Другой пример — равновесие газа в закрытом сосуде, находящемся в поле тяжести. С молекулярной точки зрения можно считать, что здесь происходят два компенсирующих друг друга процесса: вызываемый силой тяжести направленный вниз дрейф молекул компенсируется направленным вверх диффузионным потоком, который обусловлен различием концентрации молекул по высоте.

Если достаточно медленно изменять внешние условия так, чтобы при этом скорость протекающего в рассматриваемой системе процесса была значительно меньше скорости релаксации, то такой процесс будет фактически представлять собой цепочку близких друг к другу равновесных состояний. Поэтому такой процесс описывается теми же самыми макроскопическими параметрами, что и состояние

равновесия. Эти медленные процессы называются равновесными или квазистатическими. Только при таких процессах систему можно характеризовать такими параметрами, как давление, температура и т. д. Ясно, что реальные процессы являются неравновесными и могут считаться равновесными с большей или меньшей точностью.

Одной из самых важных величин, характеризующих физическую систему, является ее энергия. Понятие энергии применимо ко всем формам движения материи и объединяет самые разнообразные явления природы. Наиболее фундаментальный закон природы, установленный на основании многочисленных экспериментальных фактов и наблюдений,— это закон сохранения энергии. Согласно этому закону энергия не возникает и не исчезает: энергия неуничтожима, в явлениях природы она только переходит из одной формы в другую или от одной физической системы к другой.

Изменить энергию выделенной физической системы можно, как показывает опыт, тремя различными способами. Если внешнее воздействие на систему носит механический характер, то изменение ее энергии определяется работой внешних сил, действующих на систему. Возможно термическое внешнее воздействие, при котором изменение энергии системы происходит на молекулярном уровне без совершения макроскопической работы в результате хаотического теплового движения на границе рассматриваемой системы с другими телами. Количественной мерой изменения энергии при таком способе, называемом теплопередачей, является количество тепла, переданное системе. Наконец, изменить энергию системы можно, изменяя саму систему, например добавляя к ней или удаляя частицы, направляя на нее электромагнитное излучение и т. д. В дальнейшем мы будем рассматривать только первые два способа воздействия на систему.

Подчеркнем, что макроскопическая работа и количество тепла — это не формы энергии, а только различные способы ее изменения и передачи от одного тела к другому. Поэтому в то время как энергия характеризует состояние рассматриваемой системы, теплота и работа характеризуют изменение состояния, т. е. происходящие в системе процессы.

Полная энергия системы состоит из механической энергии и внутренней. В термодинамике обычно рассмат-

ривают покоящиеся тела, механическая энергия которых не меняется. В этом случае закон сохранения энергии можно сформулировать следующим образом: изменение внутренней энергии системы ΔU во время процесса перехода из начального состояния в конечное равно сумме совершенной над системой внешними телами работы A и полученного системой количества тепла Q:

$$\Delta U = U_2 - U_1 = A + Q. \qquad (5.1)$$

Принцип сохранения энергии применительно к термодинамическим системам (5.1) носит название первого закона термодинамики.

Процесс перехода системы из данного начального состояния в определенное конечное состояние можно осуществить различными способами. При этом совершаемая над системой работа A и передаваемое ей тепло Q зависят от способа перехода. Разумеется, их сумма, равная изменению энергии системы, во всех случаях будет одна и та же.

Рис. 5.1. p — V-диаграмма идеального газа.

Поясним это на примере простейшей термодинамической системы — идеального газа. Равновесное состояние одного моля идеального газа характеризуется заданием двух параметров, например давления и объема. Третий параметр — температура — находится при этом из уравнения состояния Менделеева — Клапейрона. Поэтому состояние моля идеального газа можно изобразить точкой на двумерной p — V-диаграмме (рис. 5.1), а квазистатический процесс, представляющий собой последовательность равновесных состояний,— непрерывной линией. Пусть точка 1 изображает начальное состояние газа, а точка 2 — конечное. Работа $\Delta A'$, совершаемая газом при изменении его объема на малую величину ΔV, равна произведению давления на изменение объема:

$$\Delta A' = p\,\Delta V. \qquad (5.2)$$

Из рис. 5.1 видно, что эта работа равна площади заштрихованной полоски высотой p и шириной ΔV. Поэтому работа,

совершаемая газом при расширении из состояния *1* в состояние *2* по пути *a*, равна площади криволинейной трапеции, ограниченной графиком процесса и осью абсцисс. Работа, совершаемая внешними силами над газом при этом процессе, на основании третьего закона Ньютона, отличается от работы газа только знаком. Теперь легко видеть, что работы внешних сил, совершаемые при расширении газа из состояния *1* в состояние *2* по путям *a* и *b*, различны. Но изменение внутренней энергии газа $U_2 - U_1$ не зависит от пути перехода. Следовательно, сообщаемые газу при этих процессах количества тепла, как следует из (5.1), будут также различны.

Получаемое системой в ходе квазистатического процесса количество тепла характеризуется физической величиной — теплоемкостью. Теплоемкостью называется отношение переданного системе на участке процесса тепла ΔQ к происшедшему на этом участке изменению температуры системы ΔT:

$$C = \frac{\Delta Q}{\Delta T}. \qquad (5.3)$$

Количество тепла, получаемое системой при изменении ее температуры на ΔT, будет неодинаковым для различных процессов. Поэтому будет разной и теплоемкость. Таким образом, теплоемкость является характеристикой не самого вещества, а определенного процесса с данным веществом.

С помощью первого закона термодинамики можно получить выражение для теплоемкости при разных процессах. Для этого удобно записать первый закон в несколько отличной от (5.1) форме:

$$Q = \Delta U + A', \qquad (5.4)$$

где A' — работа, которую совершают силы, приложенные со стороны рассматриваемой системы к внешним телам. Применяя (5.4) к малому участку процесса и рассматривая системы, у которых выражение для совершаемой работы имеет вид (5.2), получим

$$C = \frac{\Delta U}{\Delta T} + p\,\frac{\Delta V}{\Delta T}. \qquad (5.5)$$

Вычислим с помощью (5.5) теплоемкости 1 моля идеального газа при различных процессах. Внутренняя энергия

идеального газа зависит только от его температуры и для одноатомного газа, согласно (2.12), равна ${}^3/_2 RT$. Для изохорического процесса, при котором объем газа остается неизменным, $\Delta V=0$, и с помощью формулы (5.5) находим

$$C_V = \frac{\Delta U}{\Delta T} = \frac{3}{2} R. \tag{5.6}$$

Теплоемкость C_V получила название молярной теплоемкости при постоянном объеме. Она характеризует процесс нагревания газа в закрытом сосуде. А если газ заключен в цилиндр с вертикальными стенками, закрытый сверху поршнем с грузом, то нагревание газа будет происходить при постоянном давлении. Такое изобарическое нагревание сопровождается расширением газа, при котором он, поднимая поршень с грузом, совершает работу. Поэтому теплоемкость газа при постоянном давлении C_p больше, чем теплоемкость при постоянном объеме C_V. Из уравнения состояния одного моля идеального газа $pV_\mu = RT$ при постоянном давлении p находим $p\,\Delta V = R\,\Delta T$. Теперь с помощью (5.5) получаем

$$C_p = C_V + R. \tag{5.7}$$

Для одноатомного газа $C_V = {}^3/_2 R$, поэтому $C_p = {}^5/_2 R$.

Теплоемкости C_V и C_p, как видно из формул (5.6) и (5.7), постоянны. Они имеют одно и то же значение для процессов, происходящих при любых объемах и давлениях, и остаются неизменными во время процесса. Процессы, при которых теплоемкость постоянна, носят название политропических. Изохорический и изобарический процессы являются политропическими. Другим важным примером политропического процесса является адиабатический, при котором система теплоизолирована. При отсутствии теплообмена с окружающими телами первый закон термодинамики (5.1) принимает вид

$$\Delta U = A, \tag{5.8}$$

т. е. изменение внутренней энергии системы происходит только за счет работы внешних сил. Это же самое можно выразить и иначе: при адиабатическом процессе совершаемая системой работа A', согласно (5.4), равна убыли

внутренней энергии системы:

$$A' = -\Delta U = U_1 - U_2. \tag{5.9}$$

При расширении система совершает положительную работу $A' > 0$, но $A < 0$, и ее внутренняя энергия убывает. Для идеального газа, у которого внутренняя энергия зависит только от температуры, это означает, что при адиабатическом расширении газ будет охлаждаться. При адиабатическом сжатии такой газ будет нагреваться.

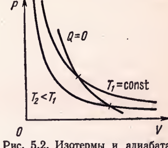


Рис. 5.2. Изотермы и адиабата на $p - V$-диаграмме идеального газа.

Такое поведение позволяет понять, почему на $p - V$-диаграмме идеального газа график обратимого адиабатического процесса идет круче графика изотермического процесса (рис. 5.2). Действительно, при изотермическом расширении температура постоянна, а при адиабатическом — убывает; поэтому адиабата должна пересекать изотермы, соответствующие все более и более низким температурам, в результате чего давление при расширении падает быстрее, чем на изотерме.

При изотермическом расширении идеальный газ совершает работу только за счет подводимого к нему тепла, так что его внутренняя энергия во время такого процесса не меняется.

Применим первый закон термодинамики к анализу работы тепловой машины. Назначение любого периодически действующего теплового двигателя состоит в совершении механической работы за счет использования внутренней энергии хаотического движения молекул, о чем кратко, но не совсем точно говорят как о превращении тепла в работу. Принципиальная схема любой тепловой машины, если отвлечься от ее конструктивных особенностей, выглядит так, как показано на рис. 5.3. Ее обязательными элементами являются два тепловых резервуара: нагреватель с некоторой температурой T_1 и холодильник с температурой T_2, меньшей температуры нагревателя. Роль холодильника может выполнять окружающая среда.

Если просто привести нагреватель в тепловой контакт с холодильником, то внутренняя энергия нагревателя будет передаваться холодильнику путем теплопередачи без совершения работы. Для совершения механической работы обязательно должно быть промежуточное звено — так называемое рабочее тело, в качестве которого может быть

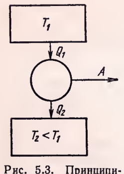


Рис. 5.3. Принципиальная схема тепловой машины.

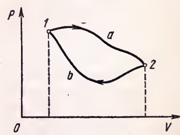


Рис. 5.4. Цикл тепловой машины на p — V-диаграмме.

использован, например, газ в цилиндре, закрытом подвижным поршнем. У периодически действующей машины все процессы с рабочим телом повторяются.

Механическая работа совершается при расширении газа в цилиндре, при котором он проходит через ряд состояний вдоль некоторой кривой a на рис. 5.4. Для того чтобы рабочее тело вернулось в исходное состояние 1, газ в цилиндре нужно сжать, для чего над ним придется совершить работу. Но эта работа обязательно должна быть меньше, чем работа, совершенная газом при расширении. На p — V-диаграмме работа газа при расширении равна площади криволинейной трапеции под кривой a. Чтобы при сжатии газа совершалась меньшая работа, кривая сжатия b должна лежать ниже кривой a: сжатие газа должно происходить при более низких температурах, чем расширение. Разность площадей под кривыми a и b, т. е. площадь, ограниченная замкнутой кривой $1 — a — 2 — b — 1$, равна работе A', совершаемой за цикл. На основании первого закона термодинамики эта работа равна разности между теплом Q_1, полученным рабочим телом за цикл

от нагревателя, и теплом Q_2, отданным им холодильнику, ибо рабочее тело после совершения цикла возвращается в исходное состояние, так что его внутренняя энергия принимает исходное значение:

$$A' = Q_1 - Q_2. \tag{5.10}$$

Отношение совершенной за цикл работы A' к количеству тепла Q_1, полученному за цикл от нагревателя, называется коэффициентом полезного действия тепловой машины η:

$$\eta = \frac{A'}{Q_1}.$$

Максимально выгодной была бы тепловая машина, у которой все полученное тепло Q_1 превращалось бы в работу: $A' = Q_1$. Однако на основании опыта установлено, что это невозможно.

К обсуждению к. п. д. тепловой машины мы еще вернемся при изучении второго закона термодинамики.

§ 6. Примеры применения первого закона термодинамики

Первый закон термодинамики, как и закон сохранения энергии в механике, часто дает возможность исследовать тепловые процессы в макроскопических системах даже в тех случаях, когда нам неизвестны детали микроскопической картины изучаемых явлений. Первый закон универсален, он применим ко всем без исключения тепловым процессам в любых системах. Как и всякий закон сохранения, он не дает детальной информации о ходе процесса, но позволяет составить уравнение баланса, если заранее известно, какие энергетические превращения происходят в рассматриваемой системе.

В этом параграфе мы рассмотрим примеры использования первого закона термодинамики.

Прежде всего сделаем несколько замечаний о смысле входящих в уравнение первого закона величин. Количество переданного тепла было определено как мера изменения внутренней энергии системы при теплопередаче. Но не всегда подведение к системе тепла приводит к изменению ее внутренней энергии. Например, при изотермическом расширении идеального газа подведение тепла не сопро-

вождается увелнчением внутренней энергнн газа. Внутренняя энергня ндеального газа зависит только от температуры н прн изотермическом процессе не меняется, но газ совершает работу, и величина этой работы равна подводимому к системе количеству тепла. Точно так же н совершение внешними силами механической работы над системой может не сопровождаться изменением ее внутренней энергин. Еслн сжнмать ндеальный газ, принимая меры к тому, чтобы его температура при этом не увеличивалась, то внутренняя энергия газа останется без изменения, а к окружающнм телам перейдет некоторое количество тепла, равное совершенной над газом прн его сжатни работе.

Применяя первый закон термодинамнкн, нужно всегда внимательно следить за тем, к какнм нзмененням в самой системе может привестн подведение к ней тепла н совершенне работы. Поясним это на следующем примере. Представьте себе, что в комнате на некоторое время включили электрическнй нагреватель, в результате чего температура воздуха увеличилась от T_1 до T_2. Может показаться, что в результате этого внутренняя энергня воздуха в комнате увелнчнлась. Проверим, так ли это. Будем счнтать воздух ндеальным газом. В состояннн теплового равновесия средняя кннетическая энергня одной молекулы пропорцнональна абсолютной температуре, а энергня всего воздуха в комнате пропорциональна числу молекул, т. е. чнслу молей газа, находящегося в комнате. Поэтому выраженне для энергин можно запнсать в виде

$$U = C_V \frac{m}{\mu} T, \tag{6.1}$$

где m — масса воздуха, μ — его молярная масса, C_V — молярная теплоемкость при постоянном объеме.

Воспользовавшись уравнением Менделеева — Клапейрона $pV = (m/\mu)RT$, замечаем, что выраженю для внутренней энергии воздуха U можно придать вид

$$U = C_V \frac{pV}{R}. \tag{6.2}$$

Объем комнаты V не изменяется прн работе нагревателя, не изменяется н давленне воздуха в комнате p. Оно равно атмосферному, поскольку комната не герметнчна.

Тогда из формулы (6.2) видно, что внутренняя энергия воздуха в комнате не изменяется при протапливании. Сразу возникают вопросы: что происходит с потребляемой от сети электроэнергией и зачем мы вообще включаем электронагреватель? Второй вопрос, конечно, шутка: для человека имеет значение не энергия воздуха, а его температура, которая повышается при протапливании. Что касается

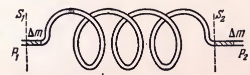


Рис. 6.1. Прохождение газа через змеевик.

энергии, потребляемой нагревателем от сети, то она целиком «выходит» наружу: масса воздуха в комнате при нагревании при постоянном давлении уменьшается.

Итак, внутренняя энергия воздуха в комнате при работе нагревателя не меняется, несмотря на повышение температуры. Но никакого парадокса в этом нет. Первый закон термодинамики, разумеется, справедлив и в этом случае. Просто при нагревании изменяется сама рассматриваемая система: число частиц в ней уменьшается.

Рассмотрим теперь стационарный поток идеального газа, протекающего через длинную спиральную трубку — змеевик (рис. 6.1). На входе змеевика поддерживаются постоянные давление p_1 и температура T_1. На выходе змеевика поддерживается постоянное давление p_2 и измеряется температура T_2. Прежде всего подсчитаем работу, совершаемую при прохождении через змеевик одного моля газа.

О какой работе идет речь? Мы знаем, что газ, а точнее, сила давления газа совершает работу при перемещении поршня или любого другого тела, ограничивающего занимаемый газом объем. Таким телом является, например, оболочка резинового шарика, деформирующаяся при его наполнении газом. Общее выражение для работы силы давления газа дается формулой (5.2):

$$\Delta A' = p\,\Delta V. \tag{6.3}$$

Отметим, что наличие тела, ограничивающего объем газа, совершенно необходимо для того, чтобы можно было говорить о совершаемой газом работе: газ, расширяющийся в пустоту, работы не совершает! Поясним это. Представим себе сосуд, разделенный перегородкой на две части: по одну сторону от перегородки находится газ, по другую — вакуум. При внезапном удалении перегородки газ заполняет весь сосуд, объем газа увеличивается ($\Delta V>0$), давление газа $p>0$, однако работы газ не совершает: $\Delta A'=0$. Формула (6.3) в этом случае неприменима.

Наряду с работой, совершаемой газом, можно рассматривать работу внешних сил, совершаемую над газом при перемещении поршня. Очевидно, что при равномерном перемещении поршня работа ΔA, совершаемая над газом, противоположна по знаку и равна по величине работе газа $\Delta A'$, т. е. $\Delta A=-p\,\Delta V$. Так о какой же работе идет речь? Ведь никаких поршней здесь нет! Однако на самом деле «поршни», т. е. внешние тела, поддерживающие заданные давления на входе и выходе змеевика, здесь есть, и они-то и совершают работу над протекающим газом. Эту работу можно подсчитать следующим образом.

Рассмотрим газ, заключенный между входным S_1 и выходным S_2 сечениями змеевика (рис. 6.1). Вследствие стационарности потока масса этого газа не меняется со временем, т. е. если за время Δt через входное сечение прошла масса газа Δm, то точно такая же масса вышла из змеевика.

Пусть объем, занимаемый массой Δm на входе при давлении p_1 и температуре T_1, равен ΔV_1, а объем на выходе из змеевика равен ΔV_2. Газ, находившийся между сечениями S_1 и S_2, теперь занимает новое положение: идущая за ним порция газа действовала на него с силой $F_1=p_1S_1$, идущий впереди газ оказывал сопротивление $F_2=p_2S_2$. Таким образом, работа внешних сил над рассматриваемым газом $\Delta A=p_1\Delta V_1-p_2\Delta V_2$. Поскольку во время процесса давления газа на входе и выходе остаются постоянными, то работа внешних сил при проталкивании через змеевик одного моля газа равна

$$A=p_1V_1-p_2V_2, \tag{6.4}$$

где V_1 и V_2 — объемы, занимаемые 1 молем газа при условиях, существующих на входе и на выходе змеевика

соответствеиио. Согласно уравнению Менделеева — Клапейрона можно написать

$$p_1V_1 = RT_1, \quad p_2V_2 = RT_2.$$

Поэтому совершаемая над одним молем газа работа равиа

$$A = R(T_1 - T_2). \tag{6.5}$$

Теперь обратим внимание на одну особенность рассматриваемого примера: в то время как температура газа на входе в змеевик T_1 задается, температура иа выходе T_2 измеряется. Напомним, что давление в обоих случаях задается. От чего же зависит температура газа иа выходе? Очевидно, от условий его прохождения через змеевик. Чем же могут различаться эти условия? Только интенсивиостью теплообмена протекающего по змеевику газа с окружающей средой. Рассмотрим вначале случай, когда теплообмеи вообще отсутствует,— змеевик адиабатически изолирован от окружающей среды. Какова при этом температура газа T_2 иа выходе? Используя первый закои термодинамики

$$Q + A = \Delta U, \tag{6.6}$$

получаем $A = \Delta U$, так как $Q = 0$. Поскольку виутренняя энергия идеального газа пропорциональна абсолютной температуре, имеем

$$\Delta U = U_2 - U_1 = C_V(T_2 - T_1). \tag{6.7}$$

Подставляя выражеиия (6.5) и (6.7) в уравнение (6.6) и учитывая, что в отсутствие теплообмена $Q = 0$, получаем

$$R(T_1 - T_2) = C_V(T_2 - T_1),$$

а так как $C_V > 0$, то отсюда немедленно следует $T_1 = T_2$.

Таким образом, температура идеального газа при адиабатическом прохождении через змеевик не меняется, а совершаемая при этом работа, как видно из (6.5), равна нулю. Для того чтобы работа была отлична от нуля, необходим теплообмен. Легко убедиться, что, когда газ получает тепло $Q > 0$, совершаемая над ним работа отрицательна и, наоборот, при $Q < 0$ $A > 0$. Действительно, подставляя в уравнение первого закона (6.6) выражения (6.5) и (6.7), получаем

$$Q = (C_V + R)(T_2 - T_1). \tag{6.8}$$

Сумма C_V+R равна молярной теплоемкости идеального газа при постоянном давлении C_p, поэтому формулу (6.8) можно записать и так:

$$Q=C_p(T_2-T_1). \qquad (6.9)$$

Из этого выражения видно, что знак $\Delta T=T_2-T_1$ совпадает со знаком Q. Если, например, $Q<0$, т. е. газ при прохождении через змеевик отдает тепло, то $T_2<T_1$, и из выражения (6.5) видно, что совершаемая при этом над газом работа положительна: $A>0$. Исходя из полученных результатов, попытаемся представить себе, как происходит протекание газа через змеевик. Если в змеевике газ охлаждается ($Q<0$), то совершаемая над газом работа положительна — внешние силы «проталкивают» газ через змеевик. Если тепло подводится ($Q>0$), то наш змеевик подобен тепловой машине — газ сам совершает работу над внешними телами. И обратите внимание, что этот результат не зависит от того, какова величина давления газа на входе и на выходе. Единственное условие при этом — давление на входе должно быть больше давления на выходе, иначе газ просто потечет в обратную сторону.

Как было выяснено, при адиабатическом протекании газа через змеевик совершаемая над ним работа равна нулю. Не кажется ли вам странным этот результат? Легко придумать такой опыт, в котором над газом работа совершается, а теплообмена с окружающей средой нет. Действительно, попробуем при помощи компрессора прокачивать через змеевик газ в вакуум. Для того чтобы процесс можно было считать стационарным, сечение выходного отверстия сделаем много меньше сечения входного. Змеевик теплоизолируем от окружающей среды.

Совершаемая компрессором над газом работа положительна и равна полной совершаемой над газом работе, ибо, как уже отмечалось, выходя в вакуум, газ работы не совершает. Так как нет обмена теплом, налицо противоречие с утверждением о том, что при адиабатическом протекании работа равна нулю.

Это противоречие возникло потому, что при прокачивании газа в вакуум происходят и такие энергетические превращения, которые были совершенно несущественны в разобранном выше примере. Действительно, первый закон термодинамики использовался в виде $Q+A=\Delta U$, где U —

внутренняя энергия газа, которая складывается из кинетической энергии хаотического движения молекул газа и потенциальной энергии их взаимодействия между собой. (Для идеального газа вторая часть — энергия взаимодействия молекул — вообще отсутствует.) Поэтому при использовании такой формулировки первого закона термодинамики заранее молчаливо предполагается, что в рассматриваемых процессах не происходит изменения механической энергии системы, т. е. не меняется потенциальная энергия газа как целого во внешнем поле, не меняется и кинетическая энергия движения газа как целого, не возникает в газе никаких макроскопических потоков. Теперь уже становится ясно, что при прокачивании газа в вакуум использовать первый закон в таком виде нельзя, так как на выходе из змеевика в вакуум возникает макроскопический направленный поток, кинетическую энергию которого необходимо учитывать. Работа компрессора в этом случае как раз и определяет кинетическую энергию этого потока.

Если вход и выход змеевика расположены на разной высоте, то в уравнении баланса энергии необходимо учитывать и изменение потенциальной энергии газа в поле тяжести, подобно тому как это делалось в гидродинамике при выводе уравнения Бернулли.

§ 7. Второй закон термодинамики. Направление тепловых процессов

Первый закон термодинамики — один из самых общих и фундаментальных законов природы. Не известно ни одного процесса, где хоть в какой-то мере наблюдалось бы его нарушение. Если какой-либо процесс запрещен первым законом, то можно быть абсолютно уверенным в том, что он никогда не произойдет. Однако этот закон не дает никаких указаний о том, в каком направлении развиваются процессы, удовлетворяющие принципу сохранения энергии.

Поясним это примерами. Первый закон термодинамики ничего не говорит о том, в каком направлении происходит теплообмен между приведенными в тепловой контакт телами, находящимися при разных температурах. Как уже обсуждалось выше, теплообмен происходит так, что температуры выравниваются и вся система стремится к состоянию теплового равновесия. Но первый закон не был бы

нарушен, если бы, наоборот, передача тепла происходила от тела с низкой температурой к телу с более высокой при условии, что полный запас внутренней энергии оставался бы неизменным. Однако повседневный опыт показывает, что само собой это никогда не происходит.

Другой пример: при падении камня с некоторой высоты вся кинетическая энергия его поступательного движения исчезает при ударе о землю, но при этом увеличивается внутренняя энергия самого камня и окружающих его тел, так что закон сохранения энергии, разумеется, не оказывается нарушенным. Но первому закону термодинамики не противоречил бы и обратный процесс, при котором к лежащему на земле камню перешло бы от окружающих предметов некоторое количество тепла, в результате чего камень поднялся бы на некоторую высоту. Однако никто никогда не наблюдал таких самопроизвольно подскакивающих камней.

Вдумываясь в эти и другие подобные примеры, мы приходим к выводу, что первый закон термодинамики не накладывает никаких ограничений на направление превращений энергии из одного вида в другой и на направление перехода тепла между телами, требуя только сохранения полного запаса энергии в замкнутых системах. Между тем опыт показывает, что разные виды энергии не равноценны в отношении способности превращаться в другие виды. Механическую энергию можно целиком превратить во внутреннюю энергию любого тела независимо от того, какова была его температура. Действительно, любое тело можно нагреть трением, увеличивая его внутреннюю энергию на величину, равную совершенной работе. Точно так же электрическая энергия может быть целиком превращена во внутреннюю, например, при прохождении электрического тока через сопротивление. А вот для обратных превращений внутренней энергии в другие виды существуют определенные ограничения, состоящие в том, что запас внутренней энергии ни при каких условиях не может превратиться целиком в другие виды энергии. С отмеченными особенностями энергетических превращений связано направление протекания процессов в природе. Второй закон термодинамики, отражающий направленность естественных процессов и налагающий ограничения на возможные направления энергетических превращений в мак-

роскопических системах, представляет собой, как и всякий фундаментальный закон, обобщение большого числа опытных фактов.

Исторически открытие второго закона термодинамики было связано с изучением вопроса о максимальном коэффициенте полезного действия тепловых машин, проведенным французским ученым Сади Карно. Позднее Клаузиус и

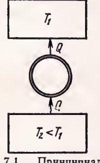


Рис. 7.1. Принципиальная схема устройства, в котором нарушается постулат Клаузиуса.

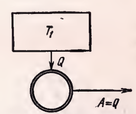


Рис. 7.2. Принципиальная схема устройства, в котором нарушается постулат Томсона.

Томсон (лорд Кельвин) предложили различные по виду, но эквивалентные формулировки второго закона термодинамики.

Согласно формулировке Клаузиуса, невозможен процесс, единственным результатом которого был бы переход тепла от тела с более низкой температурой к телу с более высокой температурой.

Томсон сформулировал второй закон термодинамики следующим образом: невозможен периодический процесс, единственным конечным результатом которого было бы совершение работы за счет тепла, взятого от одного какого-то тела.

Выражение «единственным результатом» в этих формулировках означает, что никаких других изменений, кроме указанных, ни в рассматриваемых системах, ни в окружающих их телах не происходит. Условная схема такого рода процесса, запрещенного постулатом Клаузиуса, показана на рис. 7.1, а процесса, запрещенного постулатом Томсона,— на рис. 7.2.

В формулировке Томсона второй закон термодинамики накладывает ограничения на превращение внутренней энергии в механическую. Из формулировки Томсона следует, что невозможно построить машину, которая совершала бы работу только лишь за счет получения тепла из окружающей среды. Такая гипотетическая машина получила название вечного двигателя второго рода, так как вследствие неограниченности запасов внутренней энергии в земле, океане, атмосфере такая машина была бы для всех практических целей эквивалентна вечному двигателю.

Вечный двигатель второго рода не находится в противоречии с первым законом термодинамики, в отличие от вечного двигателя первого рода, т. е. устройства для совершения работы вообще без использования источника энергии.

Эквивалентность формулировок второго закона термодинамики, предложенных Клаузиусом и Томсоном, устанавливается простыми рассуждениями.

Предположим, что постулат Томсона несправедлив. Тогда можно осуществить такой процесс, единственным результатом которого было бы совершение работы за счет тепла, взятого от единственного источника с температурой T. Эту работу можно было бы, например путем трения, снова целиком превратить в тепло, передаваемое телу, температура которого выше, чем T. Единственным результатом такого составного процесса был бы переход тепла от тела с температурой T к телу с более высокой температурой. Но это противоречило бы постулату Клаузиуса. Итак, постулат Клаузиуса не может быть справедливым, если неверен постулат Томсона.

Предположим теперь, что, наоборот, несправедлив постулат Клаузиуса, и покажем, что при этом постулат Томсона также не может выполняться. Построим обычную тепловую машину, которая будет работать, получая некоторое количество тепла Q_1 от нагревателя, отдавая Q_2 холодильнику и превращая разность $Q_1 - Q_2$ в работу (рис. 7.3).

Поскольку постулат Клаузиуса предполагается неверным, можно осуществить процесс, единственным результатом которого будет переход количества тепла, равного Q_2, от холодильника к нагревателю. Схематически это показано в правой части рис. 7.3. В результате нагре-

ватель будет отдавать рабочему телу тепловой машины тепло Q_1 и получать при процессе, противоречащем постулату Клаузиуса, тепло Q_2, так что в целом он будет отдавать количество тепла, равное Q_1-Q_2. Именно такое количество тепла машина превращает в работу. В холодильнике в целом никаких изменений вообще не происходит, ибо он отдает и получает одно и то же количество тепла Q_2. Теперь видно, что, комбинируя тепловую машину и процесс, противоречащий постулату Клаузиуса, можно получить процесс, противоречащий постулату Томсона.

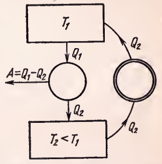


Рис. 7.3. Комбинируя с тепловой машиной устройство, изображенное на рис. 7.1, в котором нарушается постулат Клаузиуса, получаем систему, в которой нарушается постулат Томсона.

Таким образом, постулаты Клаузиуса и Томсона либо оба верны, либо оба неверны, и в этом смысле они эквивалентны. Их справедливость для макроскопических систем подтверждается всеми имеющимися экспериментальными фактами.

На основе приведенных формулировок второго закона термодинамики можно получить результаты Карно для максимально возможного коэффициента полезного действия тепловых машин. Для тепловой машины, совершающей цикл между нагревателем с фиксированной температурой T_1 и холодильником с температурой T_2, коэффициент полезного действия не может превышать значения

$$\eta=\frac{T_1-T_2}{T_1}. \tag{7.1}$$

Наибольшее значение η, определяемое формулой (7.1), достигается у тепловой машины, совершающей обратимый цикл, независимо от того, что́ используется в качестве рабочего тела. Цикл является обратимым, если он состоит из обратимых процессов, т. е. таких, которые можно провести в любом направлении через одну и ту же цепочку равновесных состояний.

Единственным обратимым циклическим процессом, который можно осуществить между нагревателем и холодиль-

ником с фиксированными температурами, является так называемый цикл Карно, состоящий из двух изотерм и двух адиабат. Для идеального газа такой цикл изображен на рис. 7.4. На участке *1—2* газ имеет температуру, равную температуре нагревателя T_1, и изотермически расширяется, получая тепло Q_1 от нагревателя. При этом газ совершает положительную работу, равную полученному теплу. На участке *2—3* газ расширяется адиабатически, и при этом его температура понижается от T_1 до значения, равного температуре холодильника T_2. Совершаемая газом на этом участке работа равна убыли его внутренней энергии. На следующем участке *3—4* газ изотермически сжимают. При этом он отдает холодильнику количество тепла Q_2, равное совершаемой над ним при сжатии работе. На участке *4—1* газ адиабатически сжимают до тех пор, пока его температура не повысится до значения T_1. Увеличение внутренней энергии газа при этом равно работе внешних сил по сжатию газа.

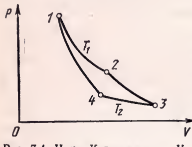


Рис. 7.4. Цикл Карно на p — V-диаграмме идеального газа.

Цикл Карно является единственным замкнутым процессом, который можно осуществить обратимым образом. В самом деле, адиабатические процессы обратимы, если их проводить достаточно медленно, т. е. квазистатически. Изотермические процессы — это единственные процессы с теплообменом, которые могут быть проведены обратимым образом. При любом другом процессе температура рабочего тела изменяется и, согласно второму закону термодинамики, теплообмен с нагревателем или холодильником не может быть обратимым: обмен теплом при наличии конечной разности температур носит характер приближения к тепловому равновесию и не является равновесным процессом.

Разумеется, обмен теплом в отсутствие разности температур происходит бесконечно медленно. Поэтому обратимый цикл Карно продолжается бесконечно долго и мощность тепловой машины при максимально возможном к. п. д.,

определяемом формулой (7.1), стремится к нулю. Процессы в любой реальной машине обязательно содержат необратимые звенья, и, следовательно, ее к. п. д. всегда меньше теоретического предела (7.1).

Преобразование внутренней энергии в механическую, как следует из второго закона термодинамики, не может быть произведено полностью. Для того чтобы превратить в механическую энергию максимально возможную часть внутренней энергии, необходимо использовать исключительно обратимые процессы. Для иллюстрации рассмотрим следующий пример. Пусть имеется некоторое тело, не находящееся в состоянии теплового равновесия с окружающей средой,

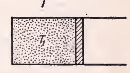


Рис. 7.5. К получению максимальной работы.

например идеальный газ в цилиндре с поршнем, имеющий температуру T_1, более высокую, чем температура окружающей среды T (рис. 7.5). Каким образом можно получить наибольшую работу при условии, что в конечном состоянии газ должен занимать тот же объем, что и в начальном?

Если бы температура газа была равна температуре окружающей среды, т. е. газ находился бы в тепловом равновесии с окружением, то никакой работы вообще получить было бы невозможно. Превращение внутренней энергии в механическую может происходить только в том случае, когда начальное состояние всей системы не является равновесным.

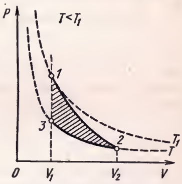


Рис. 7.6. Процесс получения максимальной работы на p — V-диаграмме.

Но и при неравновесном начальном состоянии переход системы в состояние равновесия не обязательно сопровождается превращением внутренней энергии в механическую. Если просто привести газ в тепловой контакт с окружающей средой, не давая ему расширяться, то газ

остынет и никакой работы при этом совершено не будет. Поэтому для возможности совершения работы необходимо предоставить газу возможность расширяться, имея в виду, что потом его придется сжать, так как по условию в конечном состоянии газ должен занимать тот же объём, что и в начальном. Для получения максимальной работы переход из начального состояния в конечное должен быть произведен обратимо. А это можно сделать, только используя адиабатические и изотермические процессы. Итак, газ следует адиабатически расширять до тех пор, пока его температура не станет равна температуре окружающей среды T, а затем изотермически сжать при этой температуре до исходного объема (рис. 7.6). Совершаемая газом при адиабатическом расширении *1—2* работа, как видно из рисунка, больше той работы, которую придется совершить над газом при изотермическом сжатии *2—3*. Максимальная работа, которую можно получить при переходе газа из состояния *1* в состояние *3*, равна площади заштрихованного на рис. 7.6 криволинейного треугольника *1—2—3*.

§ 8. Статистическая природа необратимости тепловых процессов

Термодинамический подход не позволяет вскрыть внутреннюю природу необратимости реальных процессов в макроскопических системах. Он только фиксирует факт необратимости во втором законе, опираясь на эксперимент. Молекулярно-кинетический подход позволяет проанализировать причины такой необратимости реальных процессов и определенной направленности энергетических превращений в природе.

Рассмотрим с точки зрения молекулярно-кинетической теории модель гипотетического «вечного» двигателя второго рода, изображенную на рис. 8.1. Предположим, что этот «вечный» двигатель работает следующим образом: газ самопроизвольно собирается в левой половине цилиндра, после чего поршень подвигают вплотную к газу. При таком перемещении внешние силы работы не совершают, так как собравшийся в левой половине газ не оказывает давления на поршень. Затем подводим к газу тепло и заставляем его изотермически расширяться до прежнего объема. При этом газ совершает работу за счет подводимого тепла. После

того, как поршень перейдет в крайнее правое положение, будем ждать, пока газ снова не соберется самопроизвольно в левой половине сосуда, и затем повторяем все снова. В результате получилась периодически действующая машина, которая совершает работу только за счет получения тепла от окружающей среды.

Молекулярно-кинетическая теория позволяет сразу объяснить, почему такое устройство не будет работать. Как мы видели, вероятность того, что газ, содержащий большое число молекул, хотя бы один раз самопроизвольно соберется в одной половине сосуда, ничтожно мала. И уж совершенно невозможно себе представить, чтобы это могло повторяться раз за разом по мере работы машины.

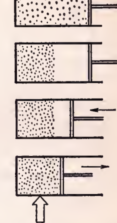


Рис. 8.1. Один из вариантов «вечного» двигателя второго рода.

Теперь можно указать, какой смысл вкладывается в понятие необратимого процесса: процесс является необратимым, если обратный процесс в действительности почти никогда не происходит. Строгого запрещения для такого процесса нет — он просто слишком маловероятен, чтобы его можно было наблюдать на опыте. Так, рассмотренный пример «вечного» двигателя второго рода основывался на процессе самопроизвольной концентрации газа в одной половине сосуда. Такой процесс является обратным для процесса расширения газа в пустоту. Расширение газа в пустоту представляет собой один из наиболее ярких примеров необратимых процессов — обратный процесс в макроскопической системе никогда не наблюдался.

Таким образом, с точки зрения представлений молекулярно-кинетической теории второй закон термодинамики утверждает то, что в природе в макроскопических системах процессы развиваются в таком направлении, когда менее вероятные состояния системы заменяются на более вероятные. Такая интерпретация второго закона термодинамики была впервые предложена Больцманом. При рассмотрении флуктуаций плотности идеального газа в § 4 было

выяснено, что состояния газа, при которых распределение молекул близко к равномерному, встречаются гораздо чаще, чем далекие от равновесия состояния с сильно неравномерным распределением молекул. Другими словами, состояния с неравномерным распределением молекул по объему, при которых число молекул в правой и левой половине сосуда сильно различается, имеют гораздо меньшую вероятность, чем состояния с почти равномерным распределением, близкие к равновесному. Итак, необратимый процесс приближения к равновесию — это переход к наиболее вероятному макроскопическому состоянию.

Сказанное выше о природе необратимости реальных процессов можно сформулировать и несколько иначе. Можно сказать, что необратимый переход к равновесию — это переход от в сильной степени упорядоченных неравновесных состояний к менее упорядоченным, хаотическим состояниям. При расширении газа в пустоту начальное состояние, когда газ занимает часть предоставленного ему объема, является в значительной мере упорядоченным, в то время как конечное состояние теплового равновесия, когда газ равномерно распределен по всему объему сосуда, является совершенно неупорядоченным. Другой пример — направленный пучок молекул газа, входящий в откачанный сосуд. Установление равновесного максвелловского распределения молекул по скоростям представляет собой необратимый процесс перехода системы из упорядоченного состояния, когда все молекулы имеют почти одинаковые по величине и направлению скорости, в конечное состояние, характеризующееся полной хаотичностью движения молекул.

С этой точки зрения легко понять устанавливаемую вторым законом термодинамики определенную направленность энергетических превращений в замкнутой макроскопической системе. Когда тело получает некоторое количество тепла за счет совершения механической работы, то это означает необратимое превращение кинетической энергии упорядоченного макроскопического движения в кинетическую энергию хаотического движения молекул. Превращение тепла в работу, наоборот, означает превращение энергии беспорядочного движения молекул в энергию упорядоченного движения макроскопического тела — такой переход, как мы видели, в принципе возможен, но исключительно маловероятен.

Необратнмый характер процессов перехода в состояиие теплового равиовесня, устанавливаемый вторым закоиом термодинамики, справедлив только для большнх макроскопнческих снстем. С термодинамической точкн зрення изолнрованиая система, пришедшая в состояние теплового равновесня, ие может самопроизвольно выйти из этого состояиня. Однако статистическая механика допускает существование флуктуаций, которые фактнчески представляют собой самопроизвольные отклонеиия снстемы от равновесия. Как уже отмечалось, чем больше частнц в системе, тем меньше относнтельная величииа флуктуацнй любого макроскопнческого параметра, н для достаточно большой системы флуктуацнями вообще можно пренебречь. Именно поэтому для такнх снстем справедлнв второй закон термодинамнкн.

В системах с иебольшнм чнслом частнц отиоснтельная величнна флуктуацнй велнка, т. е. самопронзвольные отклонения какой-лнбо величнны от ее среднего значеиня могут быть сравнимы с самим среднпм значеннем. Такая снстема часто самопроизвольио выходнт из состояння равновесня, н второй закои термоднамнкн здесь непрнменим. Характерный пример иарушеиня второго закона термодинамнкн в достаточно малых снстемах — броуновское движение, при котором взвешеиная в жндкостн частнца получает кннетнческую энергню от молекул окружающей среды, хотя температура среды ие выше, чем температура самой броуновской частицы.

ГАЗЫ, ЖИДКОСТИ, ФАЗОВЫЕ ПЕРЕХОДЫ

§ 9. Реальные газы. Уравнение Ван-дер-Ваальса

Поведение реальных газов в той нлн иной степенн отличается от поведення ндеального газа, описываемого уравнением Менделеева — Клапейроиа. Отклонення завнсят не только от того, с какнм газом — кислородом, азотом и т. д.— мы нмеем дело, но и от тех условнй, в которых находнтся газ. Чем более разрежен газ н чем выше его температура, тем менее заметны этн отклонеиня. Поэтому применнмость модели ндеального газа к определеиному реальному газу определяется не столько свойствамн самого

газа, такими, как размеры и масса его молекул и взаимодействия между ними, сколько условиями, в которых находится газ.

Взаимодействие между молекулами реального газа носит настолько сложный характер, что невозможно получить уравнение состояния, которое количественно правильно описывало бы поведение реального газа во всей области возможных изменений его температуры и плотности. Можно, однако, написать приближенное уравнение состояния реального газа, учитывающее основные качественные особенности межмолекулярного взаимодействия.

В модели идеального газа не учитывалось ни взаимодействие молекул на расстоянии, ни конечный размер самих молекул. Силы взаимодействия между молекулами имеют электромагнитную природу, хотя в целом молекулы электрически нейтральны. На больших расстояниях молекулы притягиваются друг к другу, а на малых — отталкиваются. Силы притяжения между молекулами быстро убывают с увеличением расстояния между ними, и поэтому при малой плотности газ ведет себя как идеальный. Но тем не менее силы притяжения существенны, так как именно они приводят к конденсации газа в жидкость. Отталкивание молекул на малых расстояниях настолько быстро возрастает при их сближении, что с хорошей точностью молекулы можно считать твердыми взаимно не проникающими телами и говорить об их собственном объеме. Этот объем ставит предел возможному сжатию газа.

Приступая к получению приближенного уравнения состояния реального газа, будем считать, что взаимодействие молекул приводит лишь к небольшим поправкам в уравнении состояния идеального газа. При достаточно высоких температурах и малых плотностях газа искомое уравнение должно приводить к тем же результатам, что и уравнение Менделеева — Клапейрона.

Прежде всего учтем конечные собственные размеры молекул. Фактически это приводит к тому, что предоставленный молекулам газа для движения объем будет меньше объема сосуда V. Поэтому в уравнении состояния 1 моля идеального газа заменим объем V на $V-b$, где b — характерная для данного газа положительная постоянная, учитывающая занимаемый молекулами объем:

$$p(V-b)=RT. \qquad (9.1)$$

Из этого уравнения видно, что объем газа V не может быть сделан меньше, чем b, так как при $V \to b$ давление газа p неограниченно возрастает. Разумеется, не следует ожидать, что при значениях V, близких к b, уравнение (9.1) будет правильно описывать поведение газа, ибо по своему смыслу постоянная b является малой поправкой: $b \ll V$.

Учтем теперь проявляющееся на больших расстояниях притяжение между молекулами. Притяжение должно приводить к уменьшению оказываемого газом давления на стенки сосуда, так как на каждую находящуюся вблизи стенки молекулу будет действовать со стороны остальных молекул газа сила, направленная внутрь сосуда. Поэтому давление на стенки p будет меньше значения, даваемого выражением (9.1), на некоторую величину p':

$$p = \frac{RT}{V-b} - p'. \qquad (9.2)$$

От чего может зависеть величина p'? Можно ожидать, что в грубом приближении сила, действующая на каждую молекулу со стороны остальных, будет пропорциональна числу окружающих молекул, т. е. плотности газа. Следовательно, и поправка к передаваемому отдельной молекулой импульсу при ударе о стенку пропорциональна плотности газа. Поправка же к передаваемому при ударах о стенку всеми молекулами импульсу (т. е. к давлению) будет пропорциональна квадрату плотности газа или, что то же самое, обратно пропорциональна квадрату объема. Поэтому выражение для p' можно записать в виде

$$p' = \frac{a}{V^2}, \qquad (9.3)$$

где a — характерная для данного газа положительная постоянная. Поскольку p', как и b, является малой поправкой, должно выполняться неравенство $a/V^2 \ll p$. Подставляя p' из формулы (9.3) в уравнение (9.2), получим

$$\left(p + \frac{a}{V^2}\right)(V-b) = RT. \qquad (9.4)$$

Уравнение (9.4) носит название уравнения Ван-дер-Ваальса. Оно приближенно учитывает особенности поведения реального газа, обусловленные межмолекулярным взаимодействием. Постоянные Ван-дер-Ваальса a и b определяются

опытным путем: их значения для каждого газа выбираются таким образом, чтобы уравнение (9.4) наилучшим образом описывало поведение данного газа.

Не следует рассматривать приведенные выше рассуждения как строгий вывод уравнения состояния реального газа. Они представляют собой пример феноменологического подхода, при котором качественный вид закономерности устанавливается с помощью наводящих соображений, а количественные характеристики — в данном случае постоянные a и b — находятся из сравнения с экспериментом. Как и всякое феноменологическое соотношение, уравнение Ван-дер-Ваальса в некоторой области — при достаточно высоких температурах и малых плотностях — дает правильное количественное описание свойств реального газа, тогда как во всей области изменения параметров оно дает только качественную картину поведения газа.

Кроме уравнения Ван-дер-Ваальса было предложено много других эмпирических уравнений состояния реальных газов. Некоторые из них дают лучшее согласие с опытом за счет большего числа входящих в них феноменологических постоянных. Однако при качественном исследовании поведения реальных газов удобно использовать именно уравнение Ван-дер-Ваальса благодаря его простоте и ясному физическому смыслу.

Для исследования поведения газа рассмотрим определяемые уравнением Ван-дер-Ваальса изотермы, т. е. кривые зависимости p от V при заданных значениях температуры T. С этой целью перепишем уравнение (9.4) в виде

$$V^3-\left(b+\frac{RT}{p}\right)V^2+\frac{a}{p}V-\frac{ab}{p}=0. \qquad (9.5)$$

При фиксированных значениях p и T это уравнение третьей степени относительно V. Уравнение третьей степени имеет либо один, либо три вещественных корня. Поэтому при данных значениях давления и температуры уравнение (9.5) дает либо одно, либо три значения объема. Это значит, что на p—V-диаграмме изотерма пересекает горизонтальную прямую $p=\text{const}$ либо в одной, либо в трех точках. На рис. 9.1 приведены изотермы, соответствующие различным значениям температуры $T_1>T_к>T_2>T_3$. Изотерма, соответствующая достаточно высокой температуре T_1, мало отличается от изотермы идеального газа, и одно из самых

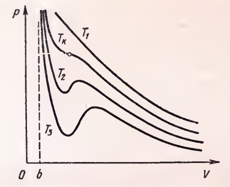


Рис. 9.1. Изотермы газа Ван-дер-Ваальса на p — V-диаграмме.

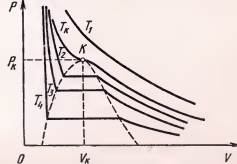


Рис. 9.2. Экспериментальные изотермы углекислого газа.

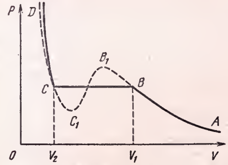


Рис. 9.3. Сравнение экспериментальной изотермы с изотермой Ван-дер-Ваальса.

существенных отличий заключается в том, что значение объема V ни при каких давлениях не может быть меньше b. При достаточно низких температурах T_2 и T_3 изотермы содержат волнообразный участок, так что прямая p=const трижды пересекает изотерму. Существует некоторая температура $T_к$, такая, что соответствующая ей изотерма отделяет все монотонные изотермы, лежащие выше нее, от «горбатых», лежащих ниже. Такая температура и соответствующая ей изотерма носят название критических.

Для того чтобы выяснить смысл этой на первый взгляд весьма странной зависимости, следует обратиться к опыту. На рис. 9.2 показаны полученные экспериментально изотермы углекислого газа CO_2. При высоких температурах эти изотермы очень похожи на изотермы идеального газа. При низких температурах характер изотерм существенно другой: у них имеется горизонтальный участок, на котором давление остается постоянным, несмотря на изменение объема.

Рассмотрим, каким физическим процессам соответствуют различные участки изотерм. При изотермическом сжатии одного моля углекислоты на участке AB (рис. 9.3) она ведет себя подобно идеальному газу: давление возрастает при уменьшении объема. При дальнейшем сжатии давление остается неизменным, а в сосуде появляется жидкая углекислота. Так продолжается на всем участке BC, пока в точке C весь газ не сконденсируется в жидкость. Участок BC соответствует равновесию между жидкой углекислотой и ее насыщенным паром. Дальнейшие попытки сжатия жидкости сопровождаются резким увеличением давления при самом незначительном уменьшении объема.

Наложим на график экспериментально полученной изотермы теоретическую кривую Ван-дер-Ваальса для той же самой температуры. На участке AB, соответствующем газообразному состоянию, обе кривые практически совпадают: уравнение Ван-дер-Ваальса вполне удовлетворительно описывает свойства паров, далеких от насыщения. Крутой ход теоретической изотермы на участке CD качественно правильно передает малую сжимаемость жидкости, хотя экспериментальная изотерма на этом участке идет еще круче. Но волнообразный участок теоретической изотермы совершенно не похож на соответствующее ему плато экспериментальной кривой.

Однако оказывается, что ряд состояний, попадающих в область «горбов», может быть экспериментально осуществлен при соблюдении определенных условий. Так, в свободном от пыли и заряженных частиц пространстве удается при постепенном сжатии получить пары при давлении, большем давления насыщенных паров при данной температуре. Такие пары называются пересыщенными, и их состояния достаточно хорошо описываются участком BB_1 теоретической изотермы. Эти состояния при появлении центров конденсации — ионов, пылинок, мельчайших капелек — становятся неустойчивыми, и пересыщенный пар мгновенно конденсируется в туман. Пересыщенный пар можно получить не только при изотермическом сжатии, но и при охлаждении насыщенного пара. Поэтому его иногда называют также переохлажденным. Такой способ получения пересыщенного пара используется в камере Вильсона, нашедшей широкое применение в ядерной физике для регистрации заряженных частиц.

Углекислота в жидком состоянии при постепенном изотермическом расширении может быть получена при давлении, меньшем давления насыщенных паров при данной температуре. Такие состояния называются перегретой жидкостью, и им соответствует участок CC_1 теоретической изотермы Ван-дер-Ваальса. Перегретая жидкость может быть получена также нагреванием жидкости при постоянном давлении, если принять меры к тому, чтобы не возникло кипения. Для этого необходимо, чтобы жидкость и стенки сосуда не содержали растворенных газов. При появлении центров кипения перегретая жидкость мгновенно закипает.

Таким образом, участки AB_1 и DC_1 теоретической изотермы соответствуют состояниям вещества, которые могут быть осуществлены экспериментально. Состояния вещества на участке C_1B_1 являются абсолютно неустойчивыми и вообще не могут осуществляться в природе, так как им соответствует уменьшение давления при сжатии. Представим на минуту, что такие состояния возможны. Пусть небольшая часть вещества случайно сжалась в результате флуктуации плотности. Тогда давление в этом месте уменьшится и станет меньше давления окружающей среды. Это приведет к дальнейшему сжатию выделенного вещества и т. д.— система самопроизвольно выходит из состояния равновесия.

Если неустойчивость состояний, соответствующих участкам CC_1 и BB_1, обуславливалась наличием центров конденсации и кипения и потому была в принципе устранимой, то неустойчивость состояний на участке C_1B_1 связана с неизбежными тепловыми флуктуациями и принципиально неустранима. Поэтому такие состояния теоретической изотермы практически никогда не осуществляются.

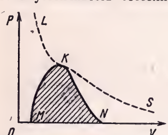


Рис. 9.4. Границы равновесия различных фаз вещества на $p - V$-диаграмме.

Наличие неосуществимого участка на изотермах Ван-дер-Ваальса при температурах ниже критической означает, что при постепенном изменении объема вещество не может все время оставаться однородным: в некоторый момент должно произойти расслоение вещества на две фазы — жидкую и газообразную.

В связи с этим интересно отметить, что эмпирическое уравнение Ван-дер-Ваальса, полученное с целью введения малых поправок к уравнению состояния идеального газа, фактически оказалось эффективным в гораздо более широкой области. В самом деле, оно указало на существование критической температуры и на необходимость расслоения вещества на фазы при температурах ниже критической. Оно также отразило возможность существования состояний пересыщенного пара и перегретой жидкости, качественно описало малую сжимаемость жидкостей и т. п.

Проследим, как изменяется по мере повышения температуры область сосуществования двух фаз на p—V-диаграмме. Прямолинейный участок экспериментальной изотермы уменьшается с ростом температуры и при критической температуре стягивается в одну точку — точку K на рис. 9.2. Если соединить между собой точки начала горизонтальных участков на всех изотермах, а также точки концов этих участков, то получим некоторую кривую MKN (рис. 9.4). Эта кривая представляет собой границу, отделяющую состояния, в которых вещество существует в двух находящихся в равновесии фазах, от однофазных состояний. Соответствующие точке K давление и объем получили

название критических. Состояние вещества в точке K называется критическим состоянием.

В критической точке сливаются в одну три точки, в которых горизонтальный участок экспериментальной изотермы пересекает изотерму Ван-дер-Ваальса.

В критическом состоянии исчезает различие между жидкостью и газом, нет и никакой границы раздела между ними.

Сжимаемость вещества в критическом состоянии неограниченно возрастает: график изотермы в окрестности этой точки идет горизонтально и, следовательно, малое изменение объема вещества не сопровождается изменением давления. Это приводит к существованию больших флуктуаций плотности в критическом состоянии, которые проявляются, например, в явлении критической опалесценции — сильном рассеянии света на случайных неоднородностях.

Границы равновесия разных фаз вещества между собой показаны на диаграмме состояний (рис. 9.4). Граница KS является чисто условной и соответствует историческому делению газообразного состояния на пар и «истинный газ» в зависимости от того, находится ли вещество при температуре ниже или выше критической. В отличие от границы MKN на этой диаграмме, при переходе через границу LKS никаких физических изменений в веществе не происходит.

§ 10. Фазовые переходы

Такие процессы, как, например, плавление твердого тела или конденсация насыщенного пара в жидкость, обычно называют фазовыми переходами.

Переход вещества из одной фазы в другую при заданном давлении происходит всегда при строго определенной температуре. Например, если нагревать лед при нормальном атмосферном давлении, то он начнет плавиться при достижении температуры 0 °C, и эта температура, несмотря на подвод тепла, будет оставаться неизменной до тех пор, пока весь лед не растает и не превратится в воду. В процессе плавления лед и вода существуют одновременно, находясь в тепловом равновесии друг с другом. Если прекратить подвод тепла, то лед и вода будут сосуществовать при этой температуре неограниченно долго.

Если изменить давление, то изменится и температура фазового перехода. Другими словами, сосуществование двух фаз в равновесии возможно лишь при строго определенном соотношении между давлением и температурой. Если изобразить эту зависимость в виде кривой в системе координат p—T, то получится так называемая диаграмма состояний.

Рассмотрим диаграмму состояний для фазового перехода между жидкостью и ее паром. Вспомним, что на p—V-диаграмме (рис. 9.2) равновесию между жидкостью и паром соответствуют горизонтальные участки экспериментальных изотерм. Горизонтальный ход изотермы на этих участках соответствует тому, что при данной температуре независимо от объема равновесие между жидкостью и паром возможно только при определенном давлении. Поэтому для получения кривой равновесия жидкости и пара на p — T-диаграмме нужно на оси абсцисс отложить значения температуры для каждой изотермы, а на оси ординат — давление, соответствующее ее горизонтальному участку. В результате получится кривая, изображенная на рис. 10.1: давление, при котором жидкость и пар могут находиться в равновесии, возрастает с увеличением температуры. Кривая сосуществования заканчивается в критической точке, так как из p—V-диаграмм видно, что по мере повышения температуры горизонтальный участок изотерм становится все меньше и исчезает при критической температуре — при температурах выше критической жидкость существовать не может. По мере приближения к критической температуре плотность насыщенного пара возрастает, при $T=T_к$ она становится равной плотности жидкости и пар вообще становится неотличимым от жидкости.

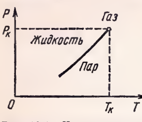
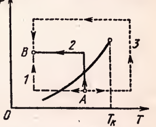


Рис. 10.1. Кривая равновесия жидкости и пара на p — T-диаграмме.

Кривая равновесия жидкости и пара на p—T-диаграмме дает зависимость давления насыщенных паров от температуры. В то же самое время эта кривая дает зависимость температуры кипения жидкости от давления, так как кипение происходит при такой температуре, когда давление насыщенного пара в пузырьках становится равным давле-

нию в жидкости. Поэтому ннога она называется крнвой кипення.

Крнвая равновесня жндкостн н пара разделяет плоскость p—T-днаграммы на две частн. Так как при заданном давленни более высоким температурам соответствует пар, а более ннзкнм температурам — жндкость и, наоборот, прн заданной температуре более низким давлениям соответствует пар, а более высоким давлениям — жидкость, то область справа н сннзу от кривой равновесия соответствует газообразной, а область слева и сверху от нее — жидкой фазе.



Рнс. 10.2. Переход нз газообразного состояння A в жидкое B по разным путям иа p — T-диаграмме.

Существованне на крнвой равновесия жндкости и пара крнтнческой точкн отчетлнво демонстрнрует отсутствне прннцнпиальной разннцы между жидкнм и газообразным состояниями вещества. Пояснны это следующнм прнмером. Возьмем пар в некотором состояннн A (рнс. 10.2) и переведем его в жидкость в состояннн, нзображаемом точкой B. Такое превращенне можно осуществить несколькнмн путямн. Можно, напрнмер, охлаждать газ при постоянном давленин p_A до тех пор, пока температура не станет равной T_B, а затем поднять давленне прн постоянной температуре до значення p_B (путь *1* на рнс. 10.2). В процессе нзобарического охлаждення путь *1* пересекает кривую равновесня, и в этой точке снстема расслаиваются на две фазы, так как происходнт конденсацня пара в жндкость. Можно перевестн снстему нз состояння A в состояние B по пути *2*, сначала нзотермнческн сжнмая газ до давления p_B, а затем изобарнчески охлаждая до температуры T_B. В этом случае конденсация пара н расслоенне на две фазы пронзойдет в процессе изотермнческого сжатия. Но переход между темн же состояннямн A и B можно провестн и такнм способом, при котором нигде не будет пронсходнть скачкообразного изменення состояння и вещество все время будет оставаться однородным. Этого можно добнться, осуществляя процесс в обход крнвой сосуществовання пара и жидкостн, например, по пути *3* на рис. 10.2. Осуществляя переход по этому

пути, мы нигде не увидели бы конденсации и, следовательно, не смогли бы сказать, что вещество перестало быть газом и стало жидкостью. И тем не менее и в этом случае мы попадем в то же самое конечное состояние, в котором рассматриваемую систему мы считаем жидкостью. В обычных условиях газы и жидкости настолько сильно отличаются друг от друга по плотности, что не представляет никакого

Рис. 10.3. Запаянная ампула с жидкостью и ее насыщенным паром.

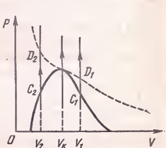


Рис. 10.4. Поведение границы раздела в зависимости от объема ампулы.

труда различить их. Но, как мы видели, различие между этими состояниями вещества в действительности не принципиальное, а скорее количественное. Различие в плотности сказывается на интенсивности взаимодействия молекул и на характере теплового движения. Но и газ при температурах выше критической можно сжать до такой степени, что его плотность станет больше характерного значения плотности жидкости.

Рассмотрим теперь процесс перехода вещества из одной фазы в другую, происходящий при неизменном объеме. Будем нагревать запаянную ампулу, содержащую некоторое количество жидкости и находящийся над ее поверхностью насыщенный пар (рис. 10.3). Если внутренний объем ампулы V_1 превышает критический объем $V_к$, соответствующий находящейся в ампуле массе вещества (рис. 10.4), то прямая, соответствующая изохорическому нагреванию, пройдет справа от критической точки и по мере нагревания количество жидкости в ампуле будет уменьшаться, пока все вещество не обратится в пар в точке C_1. При этом граница жидкости опускается и исчезает у ниж-

него конца ампулы. Если объем ампулы V_2 меньше критического объема $V_к$, то при нагревании пар будет конденсироваться, пока все вещество в ампуле не превратится в жидкость. Это произойдет в точке C_2. Граница раздела жидкости и пара поднимается и исчезает у верхнего конца ампулы. Если, наконец, объем ампулы равен критическому, то граница исчезнет где-то посредине ампулы, и произойдет это, когда температура станет равной критической. Если объем ампулы лишь немного отличается от критического, например меньше критического, то при нагревании граница жидкость — пар в ампуле перемещается вверх, но исчезает все-таки раньше, чем доходит до верхнего края ампулы. При объемах ампулы V_1 и V_2, заметно отличающихся от критического объема $V_к$, исчезновение границы раздела в точках C_1 и C_2 происходит при температурах ниже критической. При дальнейшем нагревании ампулы никаких видимых изменений вещества не происходит, в том числе и в точках D_1 и D_2 при температуре, равной критической.

Аналогичным образом на $p—T$-диаграмме состояний можно рассмотреть кривую равновесия жидкой и твердой фаз. Эта кривая выражает зависимость температуры плавления от давления и называется кривой плавления. По сравнению с кривой равновесия жидкости и пара она имеет две характерные особенности. Во-первых, на кривой плавления отсутствует критическая точка. Это связано с тем, что истинно твердые, т. е. кристаллические, тела обладают упорядоченной молекулярной структурой и принципиально отличаются от жидкостей и газов своей анизотропией. Переход между жидкостью и кристаллом не может быть произведен непрерывным образом, как это можно сделать для жидкости и газа в обход критической точки. Всегда можно точно указать, к какой фазе — жидкой или кристаллической — относится то или иное состояние вещества.

Во-вторых, наклон кривой плавления на $p—T$-диаграмме может быть разным, в отличие от кривой кипения, где давление всегда возрастает с увеличением температуры. Если при плавлении объем вещества увеличивается, то кривая плавления наклонена в ту же сторону, что и кривая кипения. Напомним, что при испарении объем всегда возрастает. Наклон кривой плавления будет противоположным, если при плавлении объем уменьшается, как это происходит, например, при таянии льда.

Две фазы вещества могут находиться в равновесии друг с другом вдоль кривых сосуществования на $p—T$-диаграмме состояний. Три фазы одного и того же вещества могут одновременно находиться в равновесии друг с другом лишь в одной определенной точке на диаграмме $p—T$. В этой точке сходятся вместе кривые равновесия каждых двух фаз, в том числе кривые плавления и кипения. Точки равновесия трех фаз называются тройными точками. Например, у воды одновременное существование в равновесии льда, воды и пара возможно только при давлении 4,62 мм рт. ст. и температуре +0,01 °C.

Рис. 10.5. $p—T$-диаграмма равновесных состояний вещества.

Тройная точка воды очень удобна в качестве стандартной точки температурной шкалы, так как ее воспроизведение, в отличие от точки кипения или плавления, не требует специального поддержания определенного давления.

Схематическая диаграмма состояний для вещества, имеющего три фазы — твердую, жидкую и газообразную,— показана на рис. 10.5. Наклон кривой плавления AB изображен для случая, когда плавление сопровождается уменьшением объема, как у льда. В отличие от кривой равновесия пара и жидкости AK, оканчивающейся при критической температуре в точке K, кривая плавления AB продолжается неограниченно. Кривая равновесия твердого тела с газом уходит в начало координат. Действительно, по законам классической механики при стремлении температуры к абсолютному нулю тепловое движение прекращается. Неподвижные атомы занимают такое расположение, при котором потенциальная энергия взаимодействия минимальна. Это расположение представляет собой регулярную пространственную решетку. Поэтому при абсолютном нуле температуры любое вещество с точки зрения классической механики должно быть кристаллическим. Существует только одно исключение из этого правила:

гелий остается жидким при всех температурах, вплоть до абсолютного нуля, если давление не превышает 25 атм.

Диаграмма состояний позволяет сразу ответить на вопрос, что произойдет с веществом при его нагревании или сжатии. Если, например, вещество в газообразном состоянии, изображаемом точкой C на рис. 10.5, подвергнуть изотермическому сжатию, то происходящий с ним процесс изобразится вертикальной пунктирной линией. Видно, что при давлении p_1 газ затвердеет, а образовавшийся кристалл при давлении p_2 расплавится.

Из диаграммы существования фаз видно, что кристалл при изобарическом нагревании не обязательно должен проходить через стадию жидкого состояния для того, чтобы превратиться в газ. Если давление выше тройной точки, то при нагревании кристалл действительно сначала расплавится, а получившаяся жидкость затем испарится. Но при давлении ниже тройной точки кристалл при нагревании сразу превращается в пар. Такой переход называется сублимацией или возгонкой. Именно так ведет себя твердая углекислота при нормальном атмосферном давлении, так как ее тройной точке соответствует давление 5,1 атм.

В заключение этого параграфа обсудим очень эффектный опыт, легко воспроизводимый с помощью самых простых средств. В стеклянный сосуд наливают некоторое количество четыреххлористого углерода (CCl_4), а сверху слой воды. При нормальном атмосферном давлении вода кипит при 100 °С, а четыреххлористый углерод — при 76,7 °С. Если медленно нагревать сосуд в водяной бане, то на границе раздела этих несмешивающихся жидкостей кипение начинается при 65,5 °С! Как объяснить это явление?

Кипение — это образование пузырьков пара внутри жидкости. Оно наступает, когда давление насыщенного пара сравняется с давлением в жидкости на той глубине, где образуется пузырек. Давление жидкости складывается из атмосферного давления и гидростатического давления столба жидкости. Если высота столба жидкости в сосуде несколько сантиметров, то гидростатическое давление составляет несколько тысячных от нормального атмосферного, и его можно не принимать во внимание. Давление насыщенных паров жидкости определяется ее температурой. В воде пузырьки содержат только пары воды, и при

100 °С давление насыщенного пара воды равно нормальному атмосферному. В четыреххлористом углероде пузырьки содержат только пары CCl_4 и давление насыщенных паров равно атмосферному при 76,7 °С.

На границе раздела этих жидкостей пузырьки содержат как пары CCl_4, так и пары воды. Давление в этих пузырьках, на основании закона Дальтона, равно сумме парциальных давлений паров воды и CCl_4. Поэтому давление, равное атмосферному, установится в пузырьках, находящихся на границе, при температуре, меньшей 76,7 °С. При этой температуре и начинается кипение на границе раздела. Опыт показывает, что это происходит при температуре 65,5 °С. Эту температуру можно установить без опыта, подбирая такую температуру по таблицам зависимости давления насыщенных паров воды и CCl_4 от температуры, при которой сумма давлений равна атмосферному. При 65,5 °С давление паров воды составляет 192 мм рт. ст., а давление паров CCl_4 — 568 мм рт. ст. Используя эти цифры, можно сделать вывод, что в каждом пузырьке, образовавшемся на границе раздела, молекул CCl_4 будет почти в три раза больше, чем молекул воды. А так как и масса молекулы четыреххлористого углерода в девять раз больше массы молекулы воды, то его испарение происходит почти в 25 раз быстрее, чем испарение воды. Наблюдая кипение в пограничном слое в течение некоторого времени, можно убедиться, что нижний слой CCl_4 выкипает значительно быстрее, чем верхний слой воды.

3. ЭЛЕКТРИЧЕСТВО И МАГНЕТИЗМ

ЭЛЕКТРОСТАТИКА

§ 1. Заряд и поле. Закон сохранения электрического заряда. Принцип суперпозиции. Теорема Гаусса

Электрический заряд и электрическое поле — первичные понятия. Бессмысленно было бы пытаться определить их через другие, более простые понятия. Все, что мы можем сделать,— это описать их свойства.

Важнейшим известным из опыта свойством электрического заряда является закон его сохранения: в изолированной системе алгебраическая сумма зарядов всех тел остается неизменной. Справедливость этого закона подтверждается не только в чисто электростатических явлениях, но и в наблюдениях над огромным числом превращений элементарных частиц. Сохранение заряда — это один из самых фундаментальных законов природы. Не известно ни одного случая его нарушения.

Заряд макроскопического тела определяется суммарным зарядом элементарных частиц, из которых состоит это тело. Зарядить макроскопическое тело можно, только изменив число содержащихся в нем заряженных элементарных частиц. Но заряд само́й элементарной частицы является ее неотъемлемым свойством, присущим ей от природы. «Зарядить» элементарную частицу, т. е. изменить ее заряд, нельзя — мы просто получим при этом другую частицу. Величина заряда всех заряженных элементарных частиц одна и та же и составляет $1{,}6\cdot10^{-19}$ Кл (или $4{,}8\cdot10^{-10}$ СГСЭ-ед. заряда). Электрический заряд частицы является ее фундаментальной характеристикой и не зависит ни от выбора системы отсчета, ни от состояния движения частицы, ни от ее взаимодействия с другими частицами.

Взаимодействие электрических зарядов, находящихся в покое, описывается законом Кулона. Этот закон устанавливает зависимость силы взаимодействия двух точечных зарядов в вакууме от их величин q_1 и q_2 и расстояния r между ними. В международной системе единиц СИ закон имеет вид

$$F=\frac{1}{4\pi\varepsilon_0}\frac{q_1q_2}{r^2}. \qquad (1.1)$$

Электрические заряды наделяют окружающее их пространство особыми физическими свойствами — создают электрическое поле. Основным свойством поля является то, что на находящуюся в этом поле заряженную частицу действует некоторая сила, т. е. взаимодействие электрических зарядов осуществляется посредством создаваемых ими полей.

Поле, создаваемое неподвижными зарядами, не меняется со временем и называется электростатическим. Силовой характеристикой электростатического поля является напряженность $\boldsymbol{E}$, измеряемая силой, действующей на единичный положительный пробный заряд. Поле, создаваемое уединенным точечным зарядом q, является сферически симметричным; величину его напряженности с помощью закона Кулона можно представить в виде

$$E=\frac{1}{4\pi\varepsilon_0}\frac{q}{r^2}. \qquad (1.2)$$

Энергетической характеристикой электростатического поля является потенциал φ, измеряемый работой, совершаемой силами поля при перемещении единичного положительного заряда из данной точки поля в некоторую точку, потенциал которой принят равным нулю. Физический смысл имеет только разность потенциалов между рассматриваемыми точками, а не абсолютные значения потенциалов точек. Существование потенциала как энергетической характеристики точки поля связано с тем, что работа сил поля при перемещении заряда не зависит от формы траектории, а определяется положением начальной и конечной точек. Поля, обладающие таким свойством, называются потенциальными.

При рассмотрении электростатического поля точечного заряда удобно в качестве точки с нулевым потенциалом

выбрать бесконечно удаленную точку. Тогда выражение для потенциала точки, отстоящей на расстоянии r от заряда, имеет вид

$$\varphi = \frac{1}{4\pi\varepsilon_0} \frac{q}{r}. \tag{1.3}$$

Потенциальная энергия W заряда q_1, помещенного в электростатическое поле, равна произведению q_1 на потенциал той точки поля, где находится этот заряд:

$$W = q_1 \varphi.$$

Если заряд q_1 находится в поле, создаваемом точечным зарядом q, то его потенциальная энергия, с учетом (1.3), равна

$$W = \frac{1}{4\pi\varepsilon_0} \frac{q q_1}{r}.$$

При одноименных зарядах q и q_1, т. е. при отталкивании, потенциальная энергия положительна. При разноименных зарядах, т. е. притяжении, потенциальная энергия, как и в гравитационном поле, отрицательна.

Электрические поля графически изображают либо с помощью силовых линий, либо с помощью эквипотенциальных поверхностей. Силовые линии перпендикулярны поверхностям постоянного потенциала; поэтому, имея одну из этих картин, мы можем построить другую.

Электрические поля удовлетворяют принципу суперпозиции: электрическое поле системы источников является суммой полей отдельных источников. Напряженность поля, создаваемого несколькими зарядами, равна векторной сумме напряженностей полей, создаваемых каждым зарядом в отдельности. Потенциал произвольной точки поля нескольких зарядов, как следует из определения потенциала, равен алгебраической сумме потенциалов, создаваемых в этой точке каждым зарядом. При этом точка нулевого потенциала выбирается общей для всех зарядов.

Принцип суперпозиции фактически означает, что присутствие других зарядов никак не сказывается на поле, создаваемом данным зарядом.

Потенциальный характер электростатического поля легко доказать для случая поля точечного заряда. Это можно сделать так же, как и для гравитационного поля, создаваемого точечной массой (см. § 8 раздела «Механика»). Тогда

из принципа суперпозиции будет следовать потенциальность любого электростатического поля, так как любое распределение электрических зарядов, создающих поле, можно представить как совокупность точечных зарядов.

Закон взаимодействия электрических зарядов — закон Кулона — можно сформулировать иначе, в виде так называемой теоремы Гаусса. Теорему Гаусса можно получить как следствие закона Кулона и принципа суперпозиции. Доказательство основывается на обратной пропорциональности силы взаимодействия двух точечных зарядов квадрату расстояния между ними. Поэтому теорема Гаусса применима к любому физическому полю, где действует закон обратных квадратов и принцип суперпозиции, например к гравитационному полю.

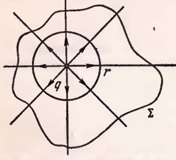


Рис. 1.1. Силовые линии электрического поля точечного заряда, пересекающие замкнутую поверхность Σ.

Прежде чем формулировать теорему Гаусса, рассмотрим картину силовых линий электрического поля неподвижного точечного заряда. Силовые линии уединенного точечного заряда представляют собой симметрично расположенные радиальные прямые (рис. 1.1). Можно провести любое число таких силовых линий. Обозначим полное их число через N. Тогда густота силовых линий на расстоянии r от заряда, т. е. число линий, пересекающих единицу поверхности сферы радиуса r, равна $N/(4\pi r^2)$. Сравнивая с выражением для напряженности поля точечного заряда (1.2), видим, что густота силовых линий и напряженность поля пропорциональны. Мы можем сделать их численно равными, надлежащим образом выбрав полное число силовых линий N:

$$\frac{N}{4\pi r^2}=\frac{q}{4\pi\varepsilon_0}\frac{1}{r^2},$$

откуда $N=q/\varepsilon_0$. Таким образом, поверхность сферы любого радиуса, охватывающей заряд q, пересекают q/ε_0 силовых линий. Так как силовые линии непрерывны, то столько же

линий выходит наружу из любой замкнутой поверхности Σ (рис. 1.1), охватывающей заряд q.

Введем теперь понятие потока вектора напряженности поля через поверхность. Произвольное поле можно мысленно разбить на малые области, в которых напряженность меняется по величине и направлению столь мало, что поле можно считать в пределах этой области однородным. В каждой такой области силовые линии представляют собой параллельные прямые и имеют постоянную густоту.

Рис. 1.2. К определению потока вектора напряженности поля через площадку ΔS.

Рассмотрим, какое число ΔN силовых линий пронизывает малую площадку ΔS, направление нормали $\boldsymbol{n}$ к которой образует угол α с направлением линий напряженности (рис. 1.2). Пусть ΔS_0 — проекция ΔS на плоскость, перпендикулярную к силовым линиям. Так как число линий, пересекающих ΔS и ΔS_0, одинаково, а густота линий, согласно принятому условию, равна величине напряженности поля E, то

$$\Delta N = E\,\Delta S_0 = E\,\Delta S \cos\alpha.$$

Величина $E\cos\alpha$ представляет собой проекцию вектора $\boldsymbol{E}$ на направление нормали $\boldsymbol{n}$ к площадке ΔS:

$$E\cos\alpha = E_n.$$

Поэтому число силовых линий ΔN, пересекающих площадку ΔS, равно

$$\Delta N = E_n\,\Delta S. \tag{1.4}$$

Произведение $E_n\Delta S$ носит название потока вектора напряженности поля через поверхность ΔS. Формула (1.4) показывает, что поток вектора $\boldsymbol{E}$ через поверхность ΔS равен числу силовых линий, пересекающих эту поверхность. Отметим, что поток вектора напряженности, как и число проходящих через поверхность силовых линий, есть скаляр. Поток вектора напряженности поля через произвольную поверхность представляет собой сумму потоков через элементарные площадки, на которые можно разбить эту поверхность. В силу соотношения (1.4) можно утверждать, что

поток напряженности поля точечного заряда q через любую охватывающую заряд замкнутую поверхность Σ (рис. 1.1), как и число выходящих из этой поверхности силовых линий, равен q/ε_0. При этом вектор нормали к элементарным площадкам замкнутой поверхности следует направлять наружу. Если заряд внутри поверхности отрицателен, то силовые линии входят внутрь этой поверхности и связанный с зарядом поток вектора напряженности поля также отрицателен.

Если внутри замкнутой поверхности находится несколько зарядов, то, в соответствии с принципом суперпозиции, будут складываться потоки напряженностей их полей. Полный поток будет равен q/ε_0, где под q следует понимать алгебраическую сумму всех зарядов, находящихся внутри поверхности.

Теперь можно окончательно сформулировать теорему Гаусса: поток N вектора напряженности электрического поля $\boldsymbol{E}$ в вакууме через любую замкнутую поверхность пропорционален полному заряду q, находящемуся внутри этой поверхности:

$$N = \frac{q}{\varepsilon_0}.$$

Теорема Гаусса широко используется в электростатике. В некоторых случаях с ее помощью легко рассчитываются поля, создаваемые симметрично распределенными зарядами.

Применим теорему Гаусса для расчета напряженности электрического поля равномерно заряженного по поверхности шара радиуса R. Распределение зарядов, создающих поле, в этом случае обладает сферической симметрией. Поэтому такой же симметрией обладает и поле. Силовые линии такого поля направлены по радиусам, а величина напряженности одинакова во всех точках, равноудаленных от центра шара. Для того чтобы найти величину напряженности поля $E(r)$ на расстоянии r от центра шара, проведем мысленно концентрическую с шаром сферическую поверхность радиуса r. Поскольку во всех точках этой сферы напряженность поля направлена перпендикулярно ее поверхности и одинакова по величине, то поток напряженности просто равен произведению величины напряженности поля на площадь поверхности сферы:

$$N = E(r)\, 4\pi r^2 \qquad (1.5)$$

Но эту величину можно выразить и с помощью теоремы Гаусса. Если нас интересует поле вне шара, т. е. при $r>R$, то $N=q/\varepsilon_0$ и, сравнивая с (1.5), находим

$$E(r)=\frac{1}{4\pi\varepsilon_0}\frac{q}{r^2} \qquad (r>R).$$

Если же интересоваться полем внутри шара, т. е. при $r<R$, то $N=0$, так как весь распределенный по поверхности шара заряд находится вне мысленно проведенной нами сферы. Поэтому поле внутри шара отсутствует:

$$E=0 \qquad (r<R).$$

Аналогично можно рассчитать электростатическое поле, создаваемое бесконечной плоскостью, заряженной с плотностью $+\sigma$, постоянной во всех точках плоскости. По соображениям симметрии можно считать, что силовые линии перпендикулярны к плоскости, направлены от нее в обе стороны и имеют всюду одинаковую густоту. Действительно, если бы густота силовых линий в разных точках была различной, то перемещение заряженной плоскости вдоль самой себя приводило бы к изменению поля в этих точках, что противоречит симметрии системы — такой сдвиг не должен изменять поле. Другими словами, поле бесконечной равномерно заряженной плоскости является однородным.

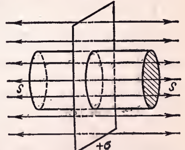


Рис. 1.3. К вычислению напряженности поля равномерно заряженной плоскости.

В качестве замкнутой поверхности для применения теоремы Гаусса выберем поверхность цилиндра, построенного следующим образом: образующая цилиндра параллельна силовым линиям, а основания имеют площади S, параллельны заряженной плоскости и лежат по разные стороны от нее (рис. 1.3). Поток напряженности поля через боковую поверхность равен нулю, поэтому полный поток через замкнутую поверхность равен сумме потоков через основания цилиндра:

$$N=2ES.$$

По теореме Гаусса этот же поток определяется зарядом той части плоскости, которая лежит внутри цилиндра, и равен $\sigma S/\varepsilon_0$. Сравнивая эти выражения для потока, находим

$$E = \frac{\sigma}{2\varepsilon_0}.$$

В качестве других примеров полей, создаваемых симметрично распределенными источниками, можно привести поле равномерно заряженной по длине бесконечной прямолинейной нити, поле равномерно заряженного бесконечного кругового цилиндра, поле шара, равномерно заряженного по объему, и т. п. Теорема Гаусса позволяет во всех этих случаях легко рассчитывать напряженность поля.

§ 2. Проводники в электрическом поле

Необходимым условием электростатического равновесия проводника является равенство нулю напряженности электрического поля внутри проводника. Если бы внутри проводника существовало макроскопическое поле, то свободные заряды (в металлах — электроны) пришли бы в движение, т. е. равновесие было бы нарушено. Условие $\boldsymbol{E}=0$ должно быть выполнено для всех точек внутри проводника независимо от того, заряжен он сам или помещен во внешнее электростатическое поле.

Условие отсутствия электростатического поля внутри проводника приводит к тому, что нескомпенсированные заряды могут располагаться только на его поверхности. В этом легко убедиться с помощью теоремы Гаусса. Рассмотрим произвольную замкнутую поверхность, ограничивающую некоторый объем внутри проводника. Во всех точках этой поверхности напряженность макроскопического электрического поля равна нулю. Следовательно, равен нулю и поток напряженности поля через эту поверхность. Тогда по теореме Гаусса равен нулю и полный заряд в объеме, ограниченном рассматриваемой поверхностью. Так как поверхность произвольна, то результат применим к любому участку внутри проводника вплоть до его границы. Итак, нескомпенсированные заряды могут располагаться только на поверхности проводника.

Отсутствие зарядов во внутренних частях проводника может быть использовано для опытной проверки закона

Кулона. Если бы в законе Кулона $F=\frac{1}{4\pi\varepsilon_0}\frac{q_1q_2}{r^2}$ показатель степени r не равнялся бы точно двум, то не была бы справедлива теорема Гаусса и во внутренних частях заряженного проводника должны были бы находиться заряды. Интересно, что отсутствие зарядов во внутренних частях заряженного металлического проводника было экспериментально установлено Кавендишем за 12 лет до того, как Кулон сформулировал закон взаимодействия точечных зарядов. Проверку закона Кулона таким способом можно произвести с большей точностью, чем при непосредственном измерении силы взаимодействия между заряженными телами, так как очень трудно создать условия, отвечающие требованию, чтобы заряды были точечными.

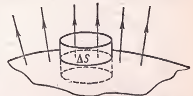


Рис. 2.1. К вычислению напряженности поля вблизи поверхности проводника.

С помощью теоремы Гаусса легко найти выражение для напряженности электрического поля в непосредственной близости от поверхности проводника. Прежде всего отметим, что во всех точках проводника потенциал одинаков и, следовательно, его граница является эквипотенциальной поверхностью, а силовые линии перпендикулярны его поверхности. Возьмем на поверхности проводника настолько малый участок ΔS, чтобы его можно было считать плоским, а поверхностную плотность заряда σ — постоянной. Проведем мысленно малую замкнутую цилиндрическую поверхность, образующие которой перпендикулярны к поверхности проводника, а основания параллельны ΔS (рис. 2.1). Нижнее основание расположено целиком внутри проводника, где поле отсутствует, а верхнее — в непосредственной близости от поверхности проводника, где силовые линии еще перпендикулярны к ней.

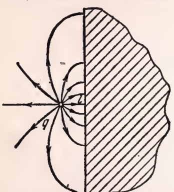
При таком выборе замкнутой поверхности поток напряженности проходит только через верхнее основание и равен $E\,\Delta S$. По теореме Гаусса

$$E\,\Delta S=\frac{\sigma\,\Delta S}{\varepsilon_0},$$

откуда

$$E = \frac{\sigma}{\varepsilon_0}. \qquad (2.1)$$

Подчеркнем, что формула (2.1) дает выражеине для напряжеиности полиого электростатнческого поля, существующего вблизи поверхиости проводника, незавнсимо от того, создается ли это поле только самнм заряжеииым



Рнс. 2.2. Электрнческое поле точечного заряда, находящегося вблнзн поверхности проводннка.

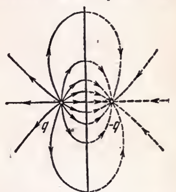


Рис. 2.3. Действне ндуцнрованных на плоской границе зарядов эквивалентно действию точечного заряда $-q$.

проводииком или еще и другими зарядамн. Из (2.1) видио, что напряженность результирующего поля вблнзн поверхностн проводнника связаиа только с плотностью зарядов на его поверхности.

В качестве примера проводника в электрическом поле рассмотрнм большой кусок металла с плоской граннцей, т. е. фактическн заполнеииое проводинком полупространство, в поле точечного заряда q, иаходящегося на расстоянии l от плоской поверхности (рис. 2.2). Выясним, какнм будет электростатическое поле во всем пространстве. Прежде всего отметим, что виутри куска металла поля нет: справа от плоскости $E=0$. Остается найти поле в полупространстве, содержащем заряд. На плоской поверхностн проводника ннduцнруются заряды, поверхиостиая плотность σ которых связана с напряжениостью полного поля вблнзи плоскостн соотношеинем (2.1). По прииципу суперпозиции полиое

поле в любой точке можно рассматривать как сумму полей заряда q и индуцированных на плоскости зарядов. Так как справа от плоскости полное поле равно нулю, то ясно, что суммарное поле всех индуцированных на плоскости зарядов можно заменить для правого полупространства полем одного точечного заряда $-q$, помещенного в ту же точку, что и исходный заряд q. Поле индуцированных зарядов симметрично относительно плоскости. Поэтому поле индуцированных зарядов в левом полупространстве эквивалентно полю точечного заряда $-q$, расположенного справа от плоскости симметрично заряду q (рис. 2.3). Итак, полное поле в левом полупространстве представляет собой суперпозицию полей, создаваемых зарядом q и зарядом $-q$, расположенным справа от плоскости симметрично заряду q.

Полученный результат можно кратко сформулировать так: действие плоской границы проводника с индуцированными на ней зарядами можно заменить действием точечного заряда $-q$, являющегося как бы зеркальным изображением данного заряда q в проводящей плоскости. Поэтому описанный способ нахождения поля носит название метода изображений.

Зная электрическое поле, можно рассчитать поверхностную плотность индуцированных на проводнике зарядов и силу, действующую на точечный заряд q. Поскольку все силовые линии, выходящие из заряда q, оканчиваются на проводящей плоскости, то полная величина индуцированного на ней заряда равна $-q$. Разумеется, этот заряд распределен неравномерно. Поверхностную плотность индуцированных зарядов легко определить с помощью соотношения (2.1). Напомним, что напряженность поля вблизи поверхности проводника направлена по нормали к ней. Очевидно, что в рассматриваемом случае поле обладает осевой симметрией: при вращении вокруг линии, соединяющей заряды q и $-q$, картина поля не меняется. Поэтому плотность заряда на поверхности зависит только от расстояния r от оси симметрии: $\sigma=\sigma(r)$. Простой расчет, идея которого понятна из рис. 2.4, приводит к результату

$$\sigma(r)=\varepsilon_0 E(r)=\frac{ql}{2\pi(r^2+l^2)^{3/2}}. \qquad (2.2)$$

Какая сила действует на заряд q? Для нахождения силы нужно знать напряженность поля, в котором нахо-

дится этот заряд. В данном случае это поле создается зарядами, индуцированными на проводнике. Точно такое же поле создавал бы заряд-изображение $-q$. Таким образом, заряд q притягивается к проводнику с такой же силой, как и к заряду $-q$, находящемуся на расстоянии $2l$ от него.

Рис. 2.4. К расчету поверхностной плотности индуцированных зарядов.

А какую работу нужно совершить, чтобы удалить заряд q на бесконечность? Может показаться, что искомая работа будет такой же, как и при раздвижении на бесконечность зарядов q и $-q$, находящихся на расстоянии $2l$ друг от друга:

$$A = \frac{1}{4\pi\varepsilon_0}\frac{q^2}{2l}. \qquad (2.3)$$

Однако это неверно! В этом можно убедиться с помощью следующего простого рассуждения. При удалении заряда q от поверхности металла будет удаляться в противоположную сторону и его «изображение» $-q$, ибо в каждый момент сила, действующая на заряд, определяется зарядом-изображением $-q$, расположенным симметрично q относительно поверхности металла. Поэтому по формуле (2.3) определяется работа, которая совершается внешними силами, действующими на оба заряда. Нам же необходимо найти работу только одной из этих сил — действующей на заряд q: ведь на самом деле никакого заряда $-q$ нет, а есть заряды, индуцированные на поверхности металла, которые при удалении заряда q растекаются по эквипотенциальной поверхности, так что при их перемещении никакой работы не совершается. Таким образом, интересующая нас работа A' будет в два раза меньше, чем работа A в (2.3):

$$A' = \frac{1}{2}A = \frac{1}{4\pi\varepsilon_0}\frac{q^2}{4l}. \qquad (2.4)$$

Этот результат становится особенно наглядным, если вспомнить, что электростатическую энергию взаимодействия зарядов можно рассматривать как энергию создаваемого ими поля. В системе двух точечных зарядов q и $-q$ поле существует во всем пространстве, в то время как в рас-

сматриваемом случае поле существует только в той половине пространства, где находится заряд q (рис. 2.3).

Мы рассмотрели простейший случай: точечный заряд вблизи бесконечной плоской поверхности проводника — и сумели просто угадать решение — заменили поле индуцированных зарядов полем фиктивного точечного заряда-изображения, расположенного по другую сторону границы

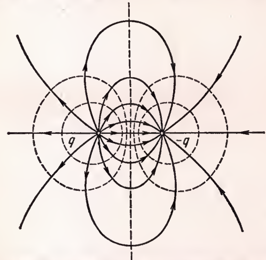


Рис. 2.5. Силовые линии и сечения эквипотенциальных поверхностей поля двух разноименных точечных зарядов.

проводника. А можно ли применять метод изображений для проводников более сложной формы? Для ответа на этот вопрос рассмотрим разобранный выше пример с несколько иной точки зрения.

Предположим, что имеются два точечных заряда q и $-q$ на расстоянии $2l$ друг от друга. Поле такой системы зарядов хорошо известно. На рис. 2.5 показаны силовые линии и сечения эквипотенциальных поверхностей. Одна из эквипотенциальных поверхностей — плоскость, перпендикулярная отрезку, соединяющему заряды, и делящая его пополам. Действительно, потенциал любой точки этой плоскости

$$\varphi = \frac{1}{4\pi\varepsilon_0}\left(\frac{q}{r_1} - \frac{q}{r_2}\right) = 0,$$

так как расстояния любой точки этой плоскости от зарядов q и $-q$ одинаковы ($r_1=r_2$). Совместим тонкий проводящий экран с этой плоскостью. Поскольку все точки проводника, помещенного в электростатическое поле, имеют одинаковый потенциал, картина поля не изменится вне экрана, а внутри него напряженность поля равна нулю. Уберем теперь заряд $-q$. Справа от экрана поля не будет, слева все останется

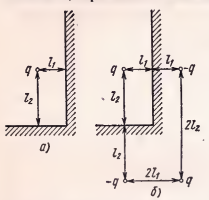


Рис. 2.6. Электрическое поле внутри двугранного угла (*а*) совпадает с полем четырех зарядов, показанных на рис. *б*.

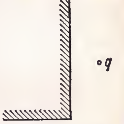


Рис. 2.7. Такую задачу методом изображений решить нельзя.

без изменения. Но получившаяся система — как раз то, что нам нужно рассмотреть! Справа от экрана поля нет, слева напряженность в любой точке определяется векторной суммой напряженностей полей, создаваемых зарядами q и $-q$, а потенциал — алгебраической суммой потенциалов этих полей.

Теперь можно сформулировать основную идею метода изображений: нужно подобрать точечные заряды, которые создавали бы такие же поля, как и индуцированные на поверхностях проводников заряды. Положение и величина этих фиктивных зарядов должны выбираться таким образом, чтобы одна из эквипотенциальных поверхностей поля, создаваемого заданными и фиктивными подобранными зарядами, совпадала бы с поверхностью проводника. Подчеркнем, что с помощью этих зарядов находится поле только вне проводников. Внутри проводников поля нет.

Метод изображений в некоторых случаях позволяет очень просто находить решения весьма сложных на пер-

вый взгляд электростатических задач. Для примера рассмотрим поле точечного заряда q, находящегося внутри проводящего прямого двугранного угла (рис. 2.6, *а*). Все электрическое поле сосредоточено только внутри угла, где расположен заряд q; по другую сторону проводящей поверхности поля нет. Нетрудно убедиться, что эквипотенциальность поверхности двугранного угла будет обеспечена, если ввести три фиктивных точечных заряда, как показано на рис. 2.6, *б*. Поэтому поле внутри угла представляет собой суперпозицию полей четырех изображенных на рис. 2.6, *б* зарядов.

Сила, с которой заряд q притягивается к проводнику, может быть представлена как векторная сумма сил его взаимодействия с тремя фиктивными зарядами.

Но, несмотря на свою привлекательность, метод изображений далеко не универсален. Достаточно поместить точечный заряд q снаружи проводящего двугранного угла (рис. 2.7), чтобы задачу уже невозможно было решить таким методом. Хотя система четырех точечных зарядов, изображенная на рис. 2.6, *б*, и в этом случае обеспечивает эквипотенциальность поверхности двугранного угла, она не дает решения задачи. Дело в том, что фиктивные заряды можно помещать только по другую от реального заряда сторону проводящей поверхности. В той точке, где находится точечный заряд, напряженность поля обращается в бесконечность. Поэтому, если мы поместим фиктивный точечный заряд по одну сторону с реальным зарядом, то в точке нахождения фиктивного заряда напряженность поля обращается в бесконечность, чего на самом деле нет.

§ 3. Конденсаторы

Под емкостью конденсатора понимают отношение заряда конденсатора к модулю разности потенциалов между его обкладками.

Что такое заряд конденсатора? Заряжая конденсатор, мы сообщаем заряды его обкладкам. При этом под зарядом обкладок нужно понимать заряды, расположенные только на внутренних, обращенных друг к другу поверхностях этих обкладок. Эти заряды равны по величине и противоположны по знаку, и абсолютная величина любого из них называется зарядом конденсатора.

Чтобы подчеркнуть различие между тем, что называют зарядом конденсатора, и полным зарядом обкладок, рассмотрим следующий пример. Пусть наружная обкладка сферического конденсатора заземлена, а внутренней сообщен заряд q. Весь этот заряд равномерно распределится по

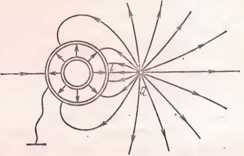


Рис. 3.1. Сферический конденсатор во внешнем электрическом поле.

внешней поверхности внутренней обкладки. Тогда на внутренней поверхности наружной сферы индуцируется заряд $-q$, и, следовательно, заряд конденсатора равен q. А что будет на внешней поверхности наружной сферы? Это зависит от того, что окружает конденсатор. Пусть, например, на расстоянии l от поверхности внешней сферы находится точечный заряд Q (рис. 3.1). Этот заряд никак не повлияет на электрическое состояние внутреннего пространства конденсатора, т. е. на поле между его обкладками. В самом деле, внутреннее и внешнее пространства разделены толщей металла наружной обкладки, в которой электрическое поле равно нулю. Но характер поля во внешнем пространстве и заряд, индуцированный на наружной поверхности внешней сферы, зависят от величины и положения заряда Q. Это поле будет точно таким же, как и в случае, когда заряд Q находится на расстоянии l от поверхности сплошного заземленного металлического шара, радиус которого равен радиусу внешней сферы конденсатора (рис. 3.2). Таким же будет и индуцированный заряд. Для нахождения его величины будем рассуждать следующим образом. Электрическое поле в любой точке пространства создается зарядом Q и зарядом, индуцированным на поверхности шара, который распределен там, разумеется, неравномерно —

как раз так, чтобы обратнлась в нуль результирующая напряженность поля внутрн шара. Согласно прннцнпу суперпознции потенцнал в любой точке можно нскать в внде суммы потенцналов полей, создаваемых точечным зарядом Q

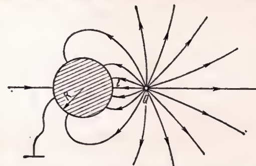


Рис. 3.2. Поле точечного заряда вблизи заземленного проводящего шара.

н точечными зарядами, на которые можно разбнть распределенный по поверхности шара индуцированный заряд. Поскольку все элементарные заряды Δq_i, на которые разбнт индуцированный на поверхностн шара заряд Q', находятся на одинаковом расстояннн R от центра шара, то потенциал создаваемого нм поля в центре шара будет равен

$$\frac{1}{4\pi\varepsilon_0}\sum_i \frac{\Delta q_i}{R}=\frac{1}{4\pi\varepsilon_0}\frac{1}{R}\sum_i \Delta q_i=\frac{1}{4\pi\varepsilon_0}\frac{Q'}{R}.$$

Тогда полный потенцнал в центре заземленного шара равен

$$\varphi_0=\frac{1}{4\pi\varepsilon_0}\left(\frac{Q'}{R}+\frac{Q}{R+l}\right)=0,$$

откуда

$$Q'=-Q\,\frac{R}{R+l}.$$

Знак мннус отражает тот факт, что нндуцнрованный заряд всегда противоположного знака.

Итак, мы вндим, что заряд на наружной поверхности внешней сферы конденсатора определяется тем окружением, в котором находится конденсатор, н не нмеет ннкакого отношення к заряду конденсатора q. Полный заряд внешней

обкладки конденсатора, разумеется, равен сумме зарядов ее внешней и внутренней поверхностей, однако заряд конденсатора определяется только зарядом внутренней поверхности этой обкладки, который связан силовыми линиями поля с зарядом внутренней обкладки.

В разобранном примере независимость электрического поля в пространстве между обкладками конденсатора и, следовательно, его емкости от внешних тел (как заряженных, так и незаряженных) обусловлена электростатической защитой, т. е. толщей металла внешней обкладки. К чему может привести отсутствие такой защиты, можно увидеть на следующем примере.

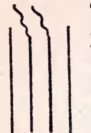


Рис. 3.3. Конденсатор в металлической коробке.

Рассмотрим плоский конденсатор в виде двух параллельных металлических пластин, электрическое поле которого практически целиком сосредоточено в пространстве между пластинами. Заключим конденсатор в незаряженную плоскую металлическую коробку, как показано на рис. 3.3. На первый взгляд может показаться, что картина поля между обкладками конденсатора не изменится, так как все поле сосредоточено между пластинами, а краевым эффектом мы пренебрегаем. Однако легко видеть, что это не так. Снаружи конденсатора поле отсутствует, поэтому во всех точках слева от конденсатора потенциал одинаков и совпадает с потенциалом левой пластины. Точно так же потенциал любой точки справа от конденсатора совпадает с потенциалом правой пластины (рис. 3.4). Поэтому, заключая конденсатор в металлическую коробку, мы соединяем проводником точки, имеющие разный потенциал. В результате в металлической коробке будет происходить перераспределение зарядов до тех пор, пока не выравняются потенциалы всех ее точек. На внутренней поверхности коробки индуцируются заряды, и появится электрическое поле внутри коробки, т. е. снаружи конденсатора (рис. 3.5). Но это означает, что на внешних поверхностях пластин конденсатора появятся заряды. Так как при этом полный заряд изолированной пластины не меняется, то заряд на ее внешней поверхности может возникнуть только за счет перетекания заряда с внутренней поверхности. Но при изменении заряда

на внутренних поверхностях обкладок изменится поле между пластинами конденсатора. Таким образом, заключение рассмотренного конденсатора в металлическую коробку приводит к изменению электрического состояния внутреннего пространства.

Изменение зарядов пластин и электрического поля в этом примере может быть легко рассчитано. Обозначим

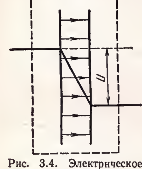


Рис. 3.4. Электрическое поле заряженного плоского конденсатора.

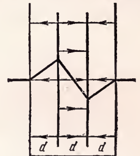


Рис. 3.5. Электрическое поле заряженного конденсатора, помещенного в металлическую коробку.

заряд изолированного конденсатора через q. Заряд, перетекший на наружные поверхности пластин при надевании коробки, обозначим через q'. Такой же заряд противоположного знака будет индуцирован на внутренних поверхностях коробки. На внутренних поверхностях пластин конденсатора останется заряд $q-q'$. Тогда в пространстве между пластинами напряженность однородного поля будет равна $(q-q')/(S\varepsilon_0)$, а вне конденсатора поле направлено в противоположную сторону и его напряженность равна $q'/(S\varepsilon_0)$, где S — площадь пластины. Требуя, чтобы разность потенциалов между противоположными стенками металлической коробки была равна нулю, и считая для простоты расстояния между всеми пластинами одинаковыми и равными d (рис. 3.5), получим

$$-\frac{q'}{S\varepsilon_0}d+\frac{q-q'}{S\varepsilon_0}d-\frac{q'}{S\varepsilon_0}d=0,$$

откуда $q'=q/3$. Это значит, что напряженность поля в конденсаторе составляет теперь 2/3 от первоначального

значення, а вие коиденсатора поле в трн раза слабее, чем было в конденсаторе до надевания коробкн. Разиость потенцналов между обкладкамн коиденсатора U_1 составляет 2/3 от первоначального напряження на конденсаторе U.

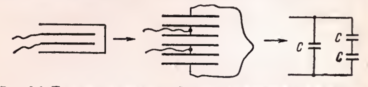
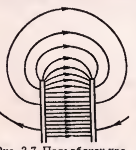


Рис. 3.6. Переход к эквивалентной схеме для конденсатора в металлической коробке.

Емкость изолнрованиого коиденсатора есть $C=q/U$. Если под емкостью снстемы, получившейся при надеваиии иа коиденсатор металлнческой коробкн, поннмать отношение полиого заряда q, сообщеииого пластине, к разиости потенциалов между пластннами U_1, то

$$C_1=\frac{q}{U_1}=\frac{3}{2}C.$$

Этот результат легко поиять, если учесть, что после надевання коробки поле существует во всех трех промежутках между пластиииами, т. е. фактически имеются три одинаковых коиденсатора, эквнвалентная схема включеиня которых показана на рис. 3.6. Вычисляя емкость получнвшейся снстемы конденсаторов, получнм $C_1={}^3/_2\,C$.



Рнс. 3.7. Поле вблизи краев пластии конденсатора.

Надетая на конденсатор металлическая коробка осуществляет электростатическую защиту системы. Теперь мы можем подиоснть снаружн к коробке любые заряженные нлн незаряженные тела н при этом электрнческое поле внутрн коробкн не изменнтся. Значнт, не измеиится н емкость системы.

Обратим виимание на то, что в разобранном прнмере, выясиив все, что иас интересовало, мы тем ие меиее обошли стороиой вопрос о том, какие же силы осуществнли перераспределеине зарядов. Какое электрическое поле вызвало движеине электронов в матернале проводящей коробки?

Очевидно, что это может быть только то поле, которое выходит за пределы конденсатора вблизи краев пластин (рис. 3.7). Хотя это поле мало́ и не принимается во внимание при расчете изменения емкости, именно оно определяет суть рассматриваемого явления — перемещает заряды и этим вызывает изменение электрического поля внутри коробки.

§ 4. Энергия электрического поля и энергия системы зарядов

Заряженный конденсатор обладает энергией. В рамках электростатики эту энергию можно рассматривать либо как потенциальную энергию взаимодействия зарядов, сосредоточенных на обкладках, либо как энергию создаваемого этими зарядами электрического поля, заключенного между обкладками конденсатора. В первом случае энергия конденсатора может быть представлена одной из следующих формул:

$$W = \frac{1}{2} qU = \frac{1}{2} CU^2 = \frac{1}{2} \frac{q^2}{C}. \qquad (4.1)$$

Энергия плоского конденсатора, рассматриваемая как энергия заключенного между его обкладками электростатического поля, имеет вид

$$W = \frac{1}{2} \varepsilon_0 E^2 V, \qquad (4.2)$$

где $V = Sd$ — объем, занимаемый полем. Коэффициент при V имеет смысл объемной плотности энергии электростатического поля.

Энергия взаимодействия системы N неподвижных точечных зарядов в вакууме выражается формулой

$$W_N = \frac{1}{2} \sum_{i=1}^{N} q_i \varphi_i, \qquad (4.3)$$

где φ_i — потенциал поля, создаваемого всеми зарядами, кроме i-го, в той точке, где находится i-й заряд:

$$\varphi_i = \frac{1}{4\pi\varepsilon_0} \sum_{\substack{k=1 \\ (k \neq i)}}^{N} \frac{q_k}{r_{ki}}. \qquad (4.4)$$

Здесь r_{ki} — расстояние между i-м и k-м зарядами.

Для доказательства этой формулы удобно использовать метод математической индукции. Прежде всего убедимся, что формула (4.3) справедлива для системы двух точечных зарядов: $N=2$. Энергия этой системы определяется работой, которая совершается внешними силами при внесении точечного заряда q_2 из бесконечности в поле, создаваемое зарядом q_1. Эта работа равна произведению заряда q_2 на потенциал φ_2 той точки поля, в которую он внесен: $W_2=q_2\varphi_2$, где $\varphi_2=q_1/(4\pi\varepsilon_0 r_{12})$. Аналогично, считая, что заряд q_1 вносится в поле, создаваемое зарядом q_2, для энергии системы получим $W_1=q_1\varphi_1$, где $\varphi_1=q_2/(4\pi\varepsilon_0 r_{12})$. Поэтому энергию системы можно записать в симметричном виде:

$$W_2=\frac{1}{2}(q_1\varphi_1+q_2\varphi_2)=\frac{1}{2}\sum_{i=1}^{2}q_i\varphi_i.$$

Теперь предположим, что формула (4.3) справедлива для N точечных зарядов, и докажем ее справедливость для системы $(N+1)$ зарядов. При внесении $(N+1)$-го заряда энергия системы изменится на величину, равную работе, совершаемой внешними силами:

$$W_{N+1}=W_N+A. \tag{4.5}$$

Здесь W_N, согласно предположению, определяется формулой (4.3), а работа, совершаемая внешними силами при перемещении заряда q_{N+1} из бесконечности в точку поля с потенциалом φ_{N+1}, равна $A=q_{N+1}\varphi_{N+1}$, где

$$\varphi_{N+1}=\frac{1}{4\pi\varepsilon_0}\sum_{i=1}^{N}\frac{q_i}{r_{i,\,N+1}} \tag{4.6}$$

— потенциал этой точки поля, создаваемый всеми зарядами, кроме q_{N+1}.

После внесения заряда q_{N+1} изменяются потенциалы всех точек поля, кроме той, где находится этот заряд. Потенциал точки, в которой находится i-й заряд, теперь будет равен φ_i':

$$\varphi_i'=\varphi_i+\frac{q_{N+1}}{4\pi\varepsilon_0 r_{i,\,N+1}}\quad (i=1,2,\ldots,N);\quad \varphi_{N+1}'=\varphi_{N+1}. \tag{4.7}$$

Теперь выразим энергию системы $(N+1)$ зарядов (4.5)

через новые значения потенциалов φ_i' с помощью соотношений (4.7):

$$W_{N+1}=\frac{1}{2}\sum_{i=1}^{N} q_i\left(\varphi_i'-\frac{q_{N+1}}{4\pi\varepsilon_0 r_{i,\,N+1}}\right)+q_{N+1}\varphi_{N+1}'.$$

Сумма произведений q_i на второе слагаемое в скобках в правой части этого равенства, в силу формулы (4.6), равна $-q_{N+1}\varphi_{N+1}'$. Поэтому

$$W_{N+1}=\frac{1}{2}\sum_{i=1}^{N} q_i\varphi_i'-\frac{1}{2}q_{N+1}\varphi_{N+1}'+q_{N+1}\varphi_{N+1}'=\frac{1}{2}\sum_{i=1}^{N+1} q_i\varphi_i'.$$

Таким образом, формула (4.3) для энергии системы точечных зарядов доказана.

Эта формула справедлива и в том случае, когда, наряду с точечными зарядами, в систему входят заряженные проводники, причем в слагаемых, соответствующих проводникам, q_i есть полный заряд проводника, а φ_i — его потенциал, создаваемый как другими зарядами, так и его собственным. Покажем это.

Пусть в системе из N зарядов есть один заряженный проводник с зарядом $q_п$ и потенциалом $\varphi_п$, а все остальные заряды — точечные. Тогда формула (4.3) должна иметь вид

$$W_N=\frac{1}{2}\sum_{i=1}^{N-1} q_i\varphi_i+\frac{1}{2}q_п\varphi_п. \qquad (4.8)$$

Разобьем мысленно заряд проводника $q_п$ на большое число M малых частей так, чтобы каждую часть Δq_k можно было считать точечным зарядом. Тогда энергию всей системы можно представить как энергию $(N-1)+M$ точечных зарядов:

$$W_N=\frac{1}{2}\sum_{i=1}^{N-1} q_i\varphi_i+\frac{1}{2}\sum_{k=1}^{M}\Delta q_k\varphi_k. \qquad (4.9)$$

Поскольку все точки проводника имеют одинаковый потенциал $\varphi_k=\varphi_п$ $(k=1,\ 2,\ldots,\ M)$, во второй сумме в этой формуле φ_k можно вынести за знак суммы:

$$\frac{1}{2}\sum_{k=1}^{M}\Delta q_k\varphi_k=\frac{1}{2}\varphi_п\sum_{k=1}^{M}\Delta q_k=\frac{1}{2}\varphi_п q_п$$

(сумма $\sum\limits_{k=1}^{M} \Delta q_k$ представляет собой полный заряд проводника $q_п$).

Таким образом, формула (4.8) доказана. Из приведенного вывода ясно, что потенциал проводника $\varphi_п$ создается как точечными зарядами q_i, так и зарядом самого проводника $q_п$. Действительно, в формуле (4.9) φ_k есть потенциал поля, создаваемого всеми зарядами, кроме Δq_k, т. е. всеми точечными зарядами q_i и зарядом проводника $q_п$, за исключением малой его части Δq_k, которая может быть выбрана сколь угодно малой по сравнению с $q_п$.

Разумеется, формула (4.3) остается справедливой и в том случае, когда в рассматриваемой системе есть только заряженные проводники и нет точечных зарядов.

Как уже отмечалось выше, электростатическую энергию можно рассматривать либо как энергию взаимодействия зарядов, либо как энергию создаваемого этими зарядами поля. Легко убедиться, что, используя формулу (4.3) для энергии взаимодействия зарядов, расположенных на обкладках конденсатора, мы немедленно получаем известную формулу энергии конденсатора:

$$W=\frac{1}{2}(q_1\varphi_1+q_2\varphi_2)=\frac{1}{2}q(\varphi_1-\varphi_2)=\frac{1}{2}qU.$$

Однако, рассматривая энергию двух разноименных точечных зарядов, мы приходим к противоречию. Согласно формуле (4.3) эта энергия отрицательна: $q_1q_2/(4\pi\varepsilon_0 r_{12})<0$, а если ее рассматривать как энергию поля этих зарядов, то энергия получается положительной, так как плотность энергии поля $^1/_2\varepsilon_0 E^2$ нигде не принимает отрицательных значений. В чем же здесь дело? Объясняется это тем, что в формуле (4.3) для энергии точечных зарядов учитывается лишь их взаимодействие, но не учитывается взаимодействие отдельных элементов каждого такого заряда между собой. Действительно, если мы имеем дело лишь с одним-единственным точечным зарядом q, то энергия, вычисляемая по формуле (4.3), равна нулю, в то время как энергия электрического поля этого заряда имеет положительное (бесконечное для истинно точечного заряда) значение, равное так называемой собственной энергии заряда q. Собственная энергия заряда может рассматриваться как энергия взаимодействия

его частей. Величина этой энергии зависит, конечно, от размеров и формы заряда. Часть ее освободилась бы при «взрыве» и разлете «осколков» заряда под действием кулоновских сил отталкивания, превратившись в кинетическую энергию «осколков», другая ее часть осталась бы в форме собственной энергии этих «осколков».

Рассмотрим теперь полную, т. е. собственную и взаимную, энергию двух зарядов q_1 и q_2. Пусть каждый из этих зарядов в отдельности создает соответственно поле $\boldsymbol{E}_1$ и $\boldsymbol{E}_2$, так что результирующее поле $\boldsymbol{E}=\boldsymbol{E}_1+\boldsymbol{E}_2$. Объемная плотность энергии поля распадается на три слагаемых в соответствии с выражением

$$E^2=(\boldsymbol{E}_1+\boldsymbol{E}_2)^2=E_1^2+E_2^2+2\,(\boldsymbol{E}_1\cdot\boldsymbol{E}_2).$$

Первые два слагаемых в правой части соответствуют объемной плотности собственных энергий зарядов q_1 и q_2, а третье слагаемое соответствует энергии взаимодействия зарядов друг с другом. Именно эта часть полной энергии системы и дается формулой (4.3). Из очевидного неравенства $(\boldsymbol{E}_1-\boldsymbol{E}_2)^2\geqslant 0$ следует, что $E_1^2+E_2^2\geqslant 2\,(\boldsymbol{E}_1\cdot\boldsymbol{E}_2)$. Таким образом, положительная собственная энергия зарядов всегда больше или в крайнем случае равна их взаимной энергии. Несмотря на то, что взаимная энергия может принимать как положительные, так и отрицательные значения, полная энергия, пропорциональная E^2, всегда положительна.

При всех возможных перемещениях зарядов, не изменяющих их формы и размеров, собственная энергия зарядов остается постоянной. Поэтому при таких перемещениях изменение полной энергии системы зарядов равно изменению их взаимной энергии. Так как во всех физических явлениях существенно именно изменение энергии системы, то постоянная часть — собственная энергия зарядов — может быть отброшена. В этом смысле и следует понимать утверждение об эквивалентности энергии взаимодействия зарядов и энергии создаваемого ими поля. Итак, мы можем сопоставлять системе зарядов либо полную энергию — энергию поля, либо энергию взаимодействия и будем получать при этом, вообще говоря, разные значения. Но, рассматривая переход системы из одного состояния в другое, мы для изменения энергии всегда получим одну и ту же величину.

Обратим внимание, что при использовании формулы (4.3) для системы точечных зарядов и проводников мы, как видно из самого вывода формулы, получаем собственную энергию проводников и взаимную потенциальную энергию всех входящих в систему зарядов, т. е. полную энергию поля за вычетом неизменной собственной энергии точечных зарядов.

Собственная энергия проводников, в отличие от собственной энергии точечных зарядов, не является постоянной. Она может измениться при изменении конфигурации системы вследствие перемещения зарядов в проводниках. Поэтому эта энергия не может быть отброшена при вычислении изменения энергии системы.

В том случае, когда система состоит только из проводников, а точечных зарядов нет, формула (4.3) дает полную энергию системы, т. е. сумму собственных энергий всех проводников и энергии их взаимодействия. Мы получаем одно и то же значение независимо от того, рассматриваем ли энергию поля или энергию системы зарядов. Примером такой системы является конденсатор, где, как мы видели, оба подхода дают одинаковый результат $W = \frac{1}{2}qU$.

В заключение этого параграфа рассмотрим энергию точечного заряда q, находящегося на расстоянии l от бесконечного плоского проводящего экрана. Поскольку потенциал проводящего экрана равен нулю, то согласно формуле (4.8) энергия системы (собственная энергия проводника и энергия взаимодействия точечного заряда с проводником) равна половине произведения заряда q на потенциал той точки, где он находится. Этот потенциал может быть найден как потенциал поля заряда-изображения, расположенного на расстоянии $2l$ от заряда q (рис. 2.3). Поэтому

$$W = \frac{1}{2} q\varphi = -\frac{1}{4\pi\varepsilon_0} \frac{q^2}{4l}$$

в полном соответствии с результатом § 2, где была вычислена работа, которую нужно совершить, чтобы удалить точечный заряд q от проводящего экрана на бесконечность (см. формулу (2.4)).

Из разобранного примера очевидно, что при наличии точечных зарядов и проводников не имеет смысла рассматривать по отдельности собственную энергию проводников и

взаимную потенциальную энергию всех зарядов, так как работа внешних сил определяет изменение суммы этих энергий. Исключить из рассмотрения можно только неизменную собственную энергию точечных зарядов.

§ 5. Энергетические превращения в конденсаторах и сохранение энергии в электростатике

Для анализа энергетических превращений, которые могут происходить в электрическом поле, рассмотрим плоский конденсатор с воздушным зазором, подсоединенный к источнику с постоянным напряжением U_0. Будем раздвигать пластины конденсатора от расстояния d_1 до расстояния d_2 в двух случаях: предварительно отсоединив конденсатор от источника питания и не отсоединяя конденсатор от источника.

В первом случае заряд на обкладках конденсатора все время остается неизменным: $q=CU=\text{const}$, хотя емкость C и напряжение U изменяются при движении пластин. Зная напряжение на конденсаторе в начальный момент, находим величину этого заряда:

$$q=C_1U_0=\frac{\varepsilon_0 S}{d_1}U_0. \quad (5.1)$$

Так как разноименно заряженные пластины конденсатора притягиваются, для их раздвижения необходимо совершить положительную механическую работу. Если при раздвижении расстояние между пластинами все время остается много меньше их линейных размеров, то, как нетрудно убедиться, сила притяжения пластин не зависит от расстояния между ними. Действительно, поле, создаваемое одной из пластин, однородно и его напряженность $E=\sigma/(2\varepsilon_0)$, где $\sigma=q/S$ — поверхностная плотность заряда. Умножив эту напряженность на величину заряда другой пластины, находим величину действующей на нее силы: $F=q^2/(2\varepsilon_0 S)$.

Для равномерного перемещения пластины внешняя сила должна уравновесить силу притяжения, и поэтому совершаемая при перемещении пластины на расстояние d_2-d_1 механическая работа равна

$$A=F(d_2-d_1)=\frac{q^2}{2\varepsilon_0 S}(d_2-d_1).$$

Подставив в эту формулу величину заряда из (5.1), найдем

$$A=\frac{\varepsilon_0 S U_0^2}{2d_1^2}(d_2-d_1). \qquad (5.2)$$

Второй случай отличается от рассмотренного тем, что при движении пластин остается неизменным не заряд конденсатора, а напряжение на нем: $U=U_0$. Поскольку расстояние между обкладками увеличивается, то напряженность поля убывает, а следовательно, убывает и заряд на пластинах. Поэтому сила притяжения пластин не остается постоянной, как в первом случае, а убывает, причем, как нетрудно убедиться, обратно пропорционально квадрату расстояния. Вычислить работу этой переменной силы методами элементарной математики затруднительно — необходимо умение интегрировать. Но все-таки мы можем найти эту работу, ибо математические трудности в физике нередко удается обойти, подходя к задаче с другой стороны, т. е. используя другие физические законы. На помощь нам придет один из самых фундаментальных законов природы — закон сохранения и превращения энергии.

Применим его сначала к более простому первому случаю. Изменение энергии W конденсатора происходит только за счет механической работы, совершаемой внешними силами: $W_2-W_1=A$. Поскольку заряд конденсатора остается неизменным, для энергии конденсатора удобно воспользоваться формулой $W=q^2/(2C)$. Таким образом,

$$A=\frac{q^2}{2}\left(\frac{1}{C_2}-\frac{1}{C_1}\right),$$

что при подстановке выражения для емкости и для заряда (5.1) приводит к окончательной формуле (5.2). Обратим внимание, что этот результат можно получить и рассматривая энергию конденсатора как энергию электрического поля между его обкладками. Так как напряженность поля и, следовательно, плотность энергии остаются неизменными, а объем, занимаемый полем, возрастает, то увеличение энергии равно произведению плотности энергии $\varepsilon_0 E^2/2$ на приращение объема $S(d_2-d_1)$.

Во втором случае энергия конденсатора изменяется как за счет механической работы, так и за счет работы, совершаемой источником питания:

$$W_2-W_1=A+A_{ист}. \qquad (5.3)$$

Определив независимо изменение энергии конденсатора и работу источника, можно с помощью закона сохранения энергии (5.3) найти механическую работу.

Поскольку в этом случае остается неизменным напряжение, для расчета энергии конденсатора удобно использовать формулу $W=CU^2/2$. Для изменения энергии получаем

$$W_2-W_1=\frac{U_0^2}{2}(C_2-C_1)=\frac{U_0^2}{2}\varepsilon_0 S\left(\frac{1}{d_2}-\frac{1}{d_1}\right). \quad (5.4)$$

При изменении заряда на обкладках конденсатора на величину $\Delta q=q_2-q_1$ источник питания совершает работу $A_{ист}=U_0(q_2-q_1)$. Заряд конденсатора определяется соотношением $q=CU_0$. Тогда

$$A_{ист}=U_0^2\varepsilon_0 S\left(\frac{1}{d_2}-\frac{1}{d_1}\right), \quad (5.5)$$

и с помощью выражения (5.3) получаем

$$A=\frac{1}{2}U_0^2\varepsilon_0 S\left(\frac{1}{d_1}-\frac{1}{d_2}\right).$$

Отметим, что из (5.5) и (5.4) видно, что

$$A_{ист}=2(W_2-W_1),$$

т. е. работа источника равна удвоенному изменению энергии конденсатора.

Интересно отметить, что как работа источника, так и изменение энергии конденсатора получились отрицательными. Это вполне понятно: совершаемая механическая работа положительна и должна была бы привести к увеличению энергии конденсатора (как и происходит в первом случае). Но энергия конденсатора убывает, и, следовательно, источник должен «принять на себя» энергию, равную убыли энергии конденсатора и механической работе внешних сил. Если процессы в источнике обратимы (аккумулятор), то он будет заряжаться, в противном случае источник просто нагревается.

Чтобы лучше разобраться в сути явлений, рассмотрим противоположный случай: присоединенные к источнику пластины конденсатора сближают от расстояния d_1 до расстояния d_2 ($d_2<d_1$). Поскольку пластины притягиваются, работа внешних сил отрицательна, ибо для равномерного

перемещения пластин внешняя сила должна быть направлена в сторону, противоположную перемещению. Энергия конденсатора при сближении пластин возрастает. Итак, механическая работа внешних сил отрицательна, а энергия конденсатора возросла, следовательно, источник совершил положительную работу. Половина этой работы равна увеличению энергии конденсатора, вторая половина передана внешним телам в виде механической работы при сближении пластин. Все приведенные выше формулы применимы, разумеется, при любом направлении перемещения пластин.

Во всех рассуждениях мы пренебрегали сопротивлением проводов, соединяющих конденсатор с источником. Если учитывать выделяющееся в проводах при движении зарядов тепло Q, уравнение баланса энергии принимает вид

$$W_2 - W_1 + Q = A + A_{\text{ист}}.$$

Изменение энергии конденсатора и работа источника выражаются, конечно, прежними формулами (5.4) и (5.5). Тепло всегда выделяется независимо от того, сближаются или раздвигаются пластины, поэтому $Q>0$. Величину Q можно вычислить, если известна скорость движения пластин. Чем больше скорость движения, тем больше выделяющееся тепло. При бесконечно медленном движении пластин $Q \to 0$.

Выше мы отметили, что работа источника питания при раздвижении пластин равна удвоенному изменению энергии конденсатора. Этот факт носит универсальный характер: если любым способом изменить энергию подсоединенного к источнику питания конденсатора, то работа, совершаемая при этом источником питания, равна удвоенному значению изменения энергии конденсатора:

$$A_{\text{ист}} = 2\,\Delta W.$$

Как в этом убедиться? Поскольку конденсатор все время остается присоединенным к источнику питания, напряжение на конденсаторе равно U_0 как в начале, так и в конце процесса (хотя во время процесса напряжение на конденсаторе может быть и меньше). Если заряд конденсатора во время процесса изменился на величину Δq, то его энергия изменилась на величину

$$\Delta W = \frac{1}{2} U_0 q_2 - \frac{1}{2} U_0 q_1 = \frac{1}{2} U_0 \Delta q.$$

При этом источник питания совершил работу

$$A_{\text{ист}} = U_0 \, \Delta q = 2\Delta W. \tag{5.6}$$

Чтобы у читателя не возникло подозрений в том, что половина энергии «бесследно исчезла», полезно написать уравнение баланса энергии:

$$A_{\text{ист}} = \Delta W + A_1 + Q, \tag{5.7}$$

где A_1 — механическая работа, совершенная при этом процессе силами, действующими на внешние тела, а Q — выделившееся тепло. Очевидно, что $A_1 + Q$ и равно оставшейся

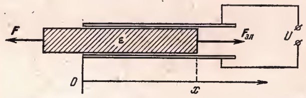


Рис. 5.1. Втягивание пластины из диэлектрика в плоский конденсатор.

половине работы источника. Существуют такие процессы, в которых либо $A_1 = 0$, либо $Q = 0$. Но, как видно из (5.6) и (5.7), изменение энергии конденсатора, соединенного с источником, обязательно сопровождается либо совершением механической работы, либо выделением тепла.

Рассмотрим теперь энергетические превращения в конденсаторах при наличии диэлектрика между обкладками. Емкость конденсатора с диэлектриком в ε раз больше, чем емкость C такого же конденсатора без диэлектрика. Конденсатор с зарядом q, отсоединенный от источника питания, обладает энергией $W = q^2/(2C)$. При заполнении пространства между обкладками диэлектриком с проницаемостью ε энергия конденсатора уменьшится в ε раз: $W' = W/\varepsilon$. Отсюда сразу можно сделать вывод о том, что диэлектрик втягивается в электрическое поле. Величина втягивающей силы при неизменном заряде конденсатора убывает по мере заполнения диэлектриком пространства между обкладками. Если на пластинах конденсатора поддерживается постоянное напряжение, то сила, втягивающая диэлектрик, не зависит от длины втянутой части. Для нахождения силы, действующей на диэлектрик со стороны электрического поля, рассмотрим втягивание твердого диэлектрика в гори-

зонтально расположенный плоский конденсатор, соединенный с источником постоянного напряжения U (рис. 5.1). Пусть под действием интересующей нас втягивающей силы $\boldsymbol{F}_{эл}$ и какой-то внешней силы $\boldsymbol{F}$ кусок диэлектрика находится в равновесии и длина втянутой части при этом равна x. Допустим, что диэлектрик вдвинулся на величину Δx. Из закона сохранения энергии следует, что совершенная источником работа $\Delta A_{ист}$ равна сумме изменения энергии конденсатора ΔW и механической работы, совершаемой силой $F_{эл}$ над внешними телами:

$$\Delta A_{ист} = \Delta W + F_{эл}\Delta x. \tag{5.8}$$

Как мы уже знаем, $\Delta A_{ист} = 2\,\Delta W$, поэтому уравнение (5.8) можно переписать в виде

$$\Delta W = F_{эл}\,\Delta x. \tag{5.9}$$

Изменение энергии конденсатора при вдвигании диэлектрика на величину Δx равно

$$\Delta W = \frac{1}{2} U^2 \Delta C = \frac{\varepsilon_0 (\varepsilon - 1)\, l\, \Delta x}{d}\, \frac{U^2}{2}, \tag{5.10}$$

где l — поперечный размер пластины.

С помощью формул (5.9) и (5.10) находим

$$F_{эл} = \frac{\Delta W}{\Delta x} = \frac{\varepsilon_0 (\varepsilon - 1)\, l}{d}\, \frac{U^2}{2},$$

т. е. сила при вдвигании диэлектрика постоянна при неизменном напряжении.

В качестве последнего примера рассмотрим задачу о втягивании жидкого диэлектрика в пространство между вертикальными пластинами плоского конденсатора, соединенного с источником постоянного напряжения U (рис. 5.2). Определим, на какой высоте h установится уровень жидкости между пластинами при погружении их концов в жидкий диэлектрик с проницаемостью ε и плотностью ρ и сколько при этом выделится тепла.

В состоянии равновесия сила, втягивающая диэлектрик в пространство между пластинами, уравновешивается силой тяжести G, действующей на поднятую жидкость: $G = \rho V g = \rho d l h g$. Для нахождения высоты подъема жидкого диэлектрика приравняем вычисленную втягивающую силу

весу столба поднявшейся жидкости и получим

$$h=\frac{\varepsilon_0(\varepsilon-1)U^2}{2\rho g d^2}. \qquad (5.11)$$

Для нахождения выделившегося при подъеме жидкости тепла проще всего исходить из закона сохранения энергии.

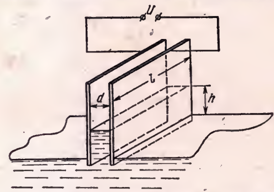


Рис. 5.2. Втягивание жидкого диэлектрика в плоский конденсатор.

Совершенная источником питания работа равна сумме изменения энергии конденсатора и механической работы, совершенной электрическими силами:

$$A_{ист}=\Delta W+A.$$

Поскольку поднятый столб жидкости покоится, совершенная механическая работа равна сумме изменения потенциальной энергии диэлектрика в поле тяжести и выделившегося тепла Q:

$$A_{ист}=\Delta W+\frac{1}{2}mgh+Q.$$

Учитывая, что $A_{ист}=2\,\Delta W$, и пользуясь соотношением (5.11), находим

$$Q=\frac{\varepsilon_0^2(\varepsilon-1)^2 lU^4}{8\rho g d^3}=\frac{1}{2}mgh.$$

Таким образом, половина механической работы, совершенной силами электрического поля, пошла на увеличение потенциальной энергии жидкого диэлектрика, а вторая поло-

вина превратилась в тепло. Как это тепло выделилось? При погружении пластин конденсатора в диэлектрик жидкость начинает подниматься, приобретая кинетическую энергию, и по инерции проскакивает положение равновесия. Возникают колебания, которые постепенно затухают из-за вязкости жидкости, и кинетическая энергия превращается в тепло. Если вязкость достаточно велика, то колебаний может и не быть — все тепло выделяется при подъеме жидкости до положения равновесия.

ПОСТОЯННЫЙ ЭЛЕКТРИЧЕСКИЙ ТОК

§ 6. Закон Ома. Работа в цепи электрического тока. Закон Джоуля — Ленца

Расчет электрических цепей постоянного тока основан на использовании закона Ома. Для участка однородной цепи (на котором не действуют сторонние силы) закон Ома выражает связь между током в цепи I, напряжением на концах участка U и сопротивлением R:

$$I=\frac{U}{R}. \tag{6.1}$$

Для неразветвленной замкнутой цепи, содержащей источник тока с электродвижущей силой $\mathscr{E}$ и внутренним сопротивлением r, закон Ома имеет вид

$$I=\frac{\mathscr{E}}{R+r}. \tag{6.2}$$

А как найти силу тока в участке неоднородной цепи, между концами которого существует некоторая разность потенциалов и внутри которого имеются скачки потенциалов, например, включен гальванический элемент или аккумулятор?

Рассмотрим для простоты участок неоднородной цепи, состоящий из двух последовательно соединенных различных проводников A и B, например медного и цинкового (рис. 6.1). Между различными проводниками имеется скачок потенциала, который не зависит от тока и существует даже в его отсутствие. Этот скачок потенциала носит название внутренней контактной разности потенциалов. Его возник-

новение обусловлено тем, что число свободных электронов в единице объема — концентрация электронного газа — различно в разных металлах. При соприкосновении таких металлов происходит диффузия электронов через контакт из того металла, где концентрация выше, в тот, где концентрация ниже. В результате между проводниками возникает разность потенциалов, величина которой определяется тем, что в установившемся динамическом равновесии диффузионный поток электронов уравновешивается встречным потоком, создаваемым возникшим в контактном слое электрическим полем. Величина скачка зависит от рода металлов и от температуры. Подчеркнем, что каждый металл остается эквипотенциальным, а скачок потенциала и связанное с ним электрическое поле имеются только в месте контакта.

Рис. 6.1. Участок неоднородной цепи.

Подсоединим теперь внешние концы проводников A и B к источнику постоянного напряжения. Обозначим потенциал левого конца проводника A через φ_1, а потенциал правого конца проводника B через φ_2 (рис. 6.1). Потенциалы металлов A и B в месте контакта обозначим через φ_A и φ_B. Так как теперь в проводниках идет ток, то, разумеется, $\varphi_1 \neq \varphi_A$ и $\varphi_2 \neq \varphi_B$. Мы пока не знаем, как записать закон Ома для всего рассматриваемого участка цепи, но зато можем написать его для каждого из однородных участков A и B. Так как проводники соединены последовательно, то через них протекает один и тот же ток I. Предположим, что ток идет слева направо, как показано на рис. 6.1. Тогда

$$\begin{aligned} \varphi_1 - \varphi_A &= IR_A, \\ \varphi_B - \varphi_2 &= IR_B, \end{aligned} \tag{6.3}$$

где R_A и R_B — сопротивления участков A и B. Сложим почленно уравнения (6.3) и перегруппируем слагаемые в левой части следующим образом:

$$(\varphi_1 - \varphi_2) + (\varphi_B - \varphi_A) = I(R_A + R_B). \tag{6.4}$$

Сумма $R_A + R_B$ есть полное сопротивление R рассматриваемого участка. Разность потенциалов $\varphi_1 - \varphi_2$ представляет

собой приложенное напряжение U. Разность $\varphi_B - \varphi_A$ есть скачок потенциала в месте контакта металлов, который, как уже отмечалось, не зависит от протекающего тока и определяется только природой металлов и температурой. Значение скачка $\varphi_B - \varphi_A$ обозначим через $\mathcal{E}$. Тогда соотношение (6.4) можно переписать в виде

$$I = \frac{U + \mathcal{E}}{R}. \tag{6.5}$$

Это и есть закон Ома для участка неоднородной цепи. Отметим, что под напряжением U на рассматриваемом участке понимается разность $\varphi_1 - \varphi_2$, где φ_1 — потенциал той точки, от которой течет ток, а φ_2 — потенциал точки, к которой течет ток. Скачок потенциала в месте контакта $\mathcal{E}$ определен как $\varphi_B - \varphi_A$, т. е. знак $\mathcal{E}$ определяется тем, повышает или понижает скачок значение потенциала в цепи в направлении протекания тока: если повышает, то $\mathcal{E} > 0$, если понижает, то $\mathcal{E} < 0$. Но ведь при рассуждениях мы выбрали направление тока слева направо наугад! А если на самом деле он течет в противоположную сторону? Предположив, что ток идет справа налево, и повторяя буквально все выкладки, мы получим значение тока, отличающееся только знаком. Это означает, что, приступая к анализу участка неоднородной цепи, мы можем вообще не задумываться о том, в какую сторону идет ток на самом деле, а задавать ему направление произвольно. Выбрав направление тока, мы определяем его величину по формуле (6.5), строго соблюдая сформулированное выше правило знаков для U и $\mathcal{E}$. Если в результате ток окажется положительным, то он действительно течет в заданном нами направлении. Если же получится отрицательное значение, то в действительности ток идет в противоположную сторону, а величина его, разумеется, найдена правильно. Ниже мы подробно рассмотрим примеры использования закона Ома для участка неоднородной цепи, иллюстрирующие сформулированное правило знаков.

Скачок потенциала на границе двух металлов мы обозначили той же буквой $\mathcal{E}$, что и электродвижущую силу. Это не случайно. Скачок потенциала возникает в результате диффузии электронов, т. е. сил неэлектростатического происхождения (не кулоновских), обусловленных хаотическим движением электронов. Такие силы неэлектроста-

тического происхождения называют сторонними. Отношение работы сторонних сил по перемещению положительного заряда вдоль некоторого участка цепи к величине этого заряда носит название электродвижущей силы на данном участке. В рассмотренном случае работа сил, вызывающих диффузию, при перемещении заряда против электрического поля в контакте определяется величиной скачка потенциала. Поэтому это действительно электродвижущая сила в обычном смысле этого слова. Но, конечно, такой контакт двух различных металлов в обычных условиях не может служить источником тока. Легко убедиться, что в замкнутой цепи из разных металлов, все участки которой поддерживаются при одной и той же температуре, сумма всех скачков потенциала равна нулю и ток в цепи отсутствует. Если поддерживать контакты при разных температурах, то сумма скачков не равна нулю и представляет собой термоэлектродвижущую силу.

Закон Ома для участка неоднородной цепи (6.5) справедлив не только в случае контактной разности потенциалов, но и для сторонних сил любой природы. Неоднородность участка может быть обусловлена наличием гальванического элемента, аккумулятора, генератора постоянного тока и т. д. Если рассматриваемый участок содержит несколько э. д. с., то в формуле (6.5) под $\mathcal{E}$ нужно понимать алгебраическую сумму всех э. д. с., причем знак каждой из них определяется в соответствии со сформулированным правилом. В этом случае R представляет собой полное сопротивление участка.

Закон Ома в форме (6.5) содержит в себе в качестве частных случаев формулы (6.1) и (6.2). Для однородного участка цепи $\mathcal{E}=0$ и (6.5) превращается в (6.1). Для неразветвленной замкнутой цепи $U=0$ и формула (6.5) переходит в (6.2).

При прохождении тока в цепи электрическое поле совершает работу, которую обычно называют работой тока. Величина работы постоянного тока I за время t на участке цепи, на концах которого поддерживается напряжение U, определяется соотношением

$$A=IUt. \qquad (6.6)$$

Прохождение тока через проводник, обладающий сопротивлением, всегда сопровождается выделением тепла. Количе-

ство выделившегося за время t тепла определяется законом Джоуля — Ленца:

$$Q = I^2Rt. \qquad (6.7)$$

В случае однородного участка, когда $I = U/R$, формулы (6.6) и (6.7) совпадают, т. е. количество выделяющегося тепла равно работе тока, и работу тока можно выразить любым из трех эквивалентных способов:

$$A = IUt = I^2Rt = \frac{U^2}{R} t. \qquad (6.8)$$

В неоднородных участках цепи, где ток определяется формулой $I = (U + \mathscr{E})/R$, выделяющееся тепло не равно работе тока. Это означает, что протекание тока в таком участке сопровождается не только выделением тепла, но и другими процессами, связанными с превращением энергии.

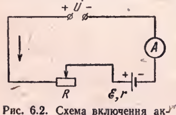


Рис. 6.2. Схема включения аккумулятора на зарядку.

В качестве примера энергетических превращений в неоднородной цепи рассмотрим зарядку аккумулятора. Не вдаваясь в детали происходящих в аккумуляторе процессов, легко сообразить, что при зарядке все химические процессы внутри него идут «вспять», и, следовательно, ток идет в направлении, противоположном току при разрядке, когда аккумулятор является источником питания для внешней цепи. Поэтому аккумулятор включается в цепь так, как показано на рис. 6.2, а ток в цепи идет в направлении, указанном стрелкой. Так как э. д. с. аккумулятора (сумма скачков потенциала внутри него) понижает потенциал в цепи в направлении протекания тока, то, в соответствии с законом Ома для неоднородного участка (6.5), ток в цепи равен

$$I = \frac{U - \mathscr{E}}{R + r}. \qquad (6.9)$$

В этой формуле r — внутреннее сопротивление аккумулятора, а сопротивление R включено в цепь для регулировки величины зарядного тока. Легко видеть, что ток будет положительным и, следовательно, пойдет в указанном

направлении только при условии, что подаваемое напряжение U больше электродвижущей силы аккумулятора $\mathscr{E}$. Только при выполнении этого условия и можно зарядить аккумулятор.

Работа, совершаемая зарядной станцией в единицу времени, т. е. работа тока на всем рассматриваемом участке, равна IU. На всех сопротивлениях, включая внутреннее сопротивление аккумулятора, в единицу времени выделяется джоулево тепло, равное $I^2(R+r)$. Кроме зарядки аккумулятора и выделения тепла, других энергетических превращений в рассматриваемой цепи не происходит. Поэтому на основании закона сохранения энергии можно утверждать, что

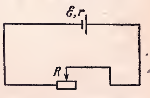


Рис. 6.3. К исследованию условий работы источника тока.

$$IU = I^2(R+r) + P_{\text{зар}}, \quad (6.10)$$

где $P_{\text{зар}}$ — мощность, идущая непосредственно на зарядку аккумулятора. Подставляя в (6.10) выражение для силы тока (6.9), получим

$$P_{\text{зар}} = \frac{\mathscr{E}(U-\mathscr{E})}{R+r} = \mathscr{E}I. \qquad (6.11)$$

Таким образом, при зарядке аккумулятор в единицу времени запасает энергию, равную $I\mathscr{E}$. Разумеется, этого результата можно было ожидать из элементарных соображений: ведь процессы в аккумуляторе считаются обратимыми, а при разрядке аккумулятор развивает мощность $I\mathscr{E}$.

Обратим внимание, что, считая известными выражения для полной работы тока, для джоулева тепла и для работы зарядки аккумулятора, можно с помощью закона сохранения энергии получить выражение (6.9) для тока в цепи, т. е. закон Ома для данного случая. Для этого нужно просто подставить в (6.10) $P_{\text{зар}}=I\mathscr{E}$.

В заключение этого параграфа исследуем условия работы источника постоянного тока, замкнутого на внешнее сопротивление R (рис. 6.3): каким должно быть сопротивление нагрузки R для того, чтобы получить максимальную силу тока в цепи, максимальную полезную мощность, максимальный коэффициент полезного действия?

Ток в цепи определяется законом Ома (6.2): $I=\mathscr{E}/(R+r)$. Полная мощность P, развиваемая источником тока, равна $I\mathscr{E}=\mathscr{E}^2/(R+r)$. Полезная мощность $P_п$, т. е. мощность, выделяющаяся на нагрузке, дается соотношением

$$P_п=I^2R=\frac{\mathscr{E}^2R}{(R+r)^2}. \qquad (6.12)$$

Коэффициент полезного действия η источника в этой цепи, определяемый как отношение полезной мощности к полной, зависит от сопротивления нагрузки:

$$\eta=\frac{P_п}{P}=\frac{R}{R+r}. \qquad (6.13)$$

Исследуем полученные выражения. Полная мощность P и ток в цепи I отличаются постоянным множителем $\mathscr{E}$, поэтому их зависимость от сопротивления нагрузки R одинакова (кривая 1 на рис. 6.4). Максимальное значение этих величин получается при $R=0$, т. е. при коротком замыкании источника. Как видно из формул (6.12) и (6.13), при этом равны нулю полезная мощность $P_п$ и коэффициент полезного действия η. При $R=r$ полная мощность и ток равны половине своего максимального значения, коэффициент полезного действия η равен 0,5, а полезная мощность достигает своего максимального значения, равного половине полной мощности P при этой нагрузке. Для того чтобы убедиться, что при равенстве сопротивления нагрузки и внутреннего сопротивления источника тока полезная мощность максимальна, преобразуем правую часть выражения (6.12) следующим образом:

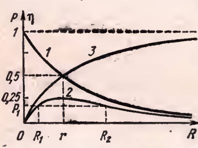


Рис. 6.4. Зависимость мощности и к. п. д. источника тока от сопротивления нагрузки.

$$P_п=\frac{\mathscr{E}^2}{(R+r)^2/R}=\frac{\mathscr{E}^2}{R+2r+r^2/R}. \qquad (6.14)$$

Полезная мощность будет максимальной, когда знаменатель правой части выражения (6.14) минимален.

Преобразуем знаменатель:

$$R-2r+\frac{r^2}{R}+4r=\left(\sqrt{R}-\frac{r}{\sqrt{R}}\right)^2+4r. \qquad (6.15)$$

Функция (6.15) достигает минимума тогда, когда выражение в скобках равно нулю, т. е. при $R=r$.

При неограниченном увеличении сопротивления нагрузки ($R\to\infty$) как полная, так и полезная мощность стремятся к нулю (кривая *2*), а коэффициент полезного действия — к единице (кривая *3*).

Из рис. 6.4 видно, что требования получения максимального тока в цепи, максимальной полезной мощности и максимального к. п. д. противоречивы. Для получения возможно бо́льшего тока сопротивление нагрузки должно быть малым по сравнению с внутренним сопротивлением источника, но при этом близки к нулю полезная мощность и к. п. д.: почти вся совершаемая источником тока работа идет на выделение тепла на внутреннем сопротивлении r. Чтобы получить от данного источника тока максимальную полезную мощность, следует взять нагрузку с сопротивлением R, равным внутреннему сопротивлению источника. Величина максимальной полезной мощности $P_{\text{п max}}=\mathscr{E}^2/(4r)$, но коэффициент полезного действия при этом равен всего лишь 0,5. Любую полезную мощность P_1, меньшую максимальной, мы можем получить, как свидетельствует ход кривой *2*, при двух значениях R_1 и R_2 сопротивления нагрузки. Практически для получения заданной полезной мощности следует выбирать нагрузку с бо́льшим сопротивлением R_2, так как к. п. д. при этом выше. Для получения к. п. д., близкого к единице, следует брать нагрузку с сопротивлением, много большим внутреннего сопротивления источника тока, но при этом выделяющаяся мощность $P_{\text{п}}\to 0$.

§ 7. Расчет цепей постоянного тока. Правила Кирхгофа

Закон Ома для неоднородного участка цепи (6.5) позволяет рассчитать любую сложную разветвленную цепь постоянного тока. Такие расчеты удобно проводить, пользуясь двумя правилами, сформулированными Кирхгофом.

Рассмотрим произвольную разветвленную цепь, часть которой изображена на рис. 7.1. Первое правило Кирхгофа относится к узлам, т. е. точкам, в которых сходится не менее трех проводников. Вследствие закона сохранения заряда в любой точке цепи, в том числе и в любом узле, при прохождении постоянного тока не должно происходить накопления электрического заряда. Поэтому сумма притекающих

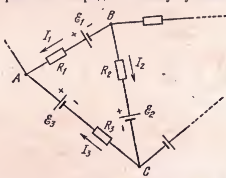


Рис. 7.1. Часть разветвленной электрической цепи.

к узлу токов должна равняться сумме утекающих. Если условиться считать подходящие к узлу токи положительными, а исходящие из узла — отрицательными, то можно сказать, что алгебраическая сумма сил токов в узле равна нулю:

$$\sum_{k=1}^{n} I_k = 0, \qquad (7.1)$$

где n обозначает число проводов, сходящихся в узле.

Второе правило Кирхгофа относится к произвольным замкнутым контурам, которые можно выделить в рассматриваемой разветвленной цепи. Рассмотрим контур $ABCA$ на рис. 7.1. Поскольку при расчете мы будем использовать закон Ома для неоднородного участка цепи, то, как мы видели, направление токов в неразветвленных участках можно задать произвольно, например так, как на рис. 7.1. (Напомним, что если в результате расчета какой-либо из токов окажется отрицательным, то это означает, что в действительности ток на этом участке течет в противоположную сторону.) Запишем закон Ома для каждого из неразвет-

вленных участков контура $ABCA$. Обозначив потенциалы узлов через φ_A, φ_B и φ_C, получим

$$\begin{aligned} I_1R_1 &= \varphi_B - \varphi_A + \mathscr{E}_1, \\ I_2R_2 &= \varphi_B - \varphi_C - \mathscr{E}_2, \\ I_3R_3 &= \varphi_C - \varphi_A + \mathscr{E}_3. \end{aligned} \tag{7.2}$$

В этих формулах через R_k обозначено полное сопротивление участка, по которому течет ток I_k. Легко заметить, что если первое уравнение системы (7.2) умножить на -1 и затем сложить почленно все три уравнения, то потенциалы узлов выпадают:

$$-I_1R_1 + I_2R_2 + I_3R_3 = -\mathscr{E}_1 - \mathscr{E}_2 + \mathscr{E}_3. \tag{7.3}$$

Глядя на формулу (7.3), нетрудно сформулировать правило, с помощью которого можно было бы непосредственно получить это равенство: нужно выбрать определенное направление обхода замкнутого контура (например, по часовой стрелке) и приравнять алгебраическую сумму произведений сил токов на сопротивления соответствующих участков алгебраической сумме э. д. с., встречающихся в этом контуре. При этом ток считается положительным, если его направление совпадает с направлением обхода контура, и отрицательным в противоположном случае; э. д. с. берется со знаком «+», если она повышает потенциал в цепи в направлении обхода контура, и со знаком «—», если понижает. Это и есть второе правило Кирхгофа, которое можно коротко записать так:

$$\sum_{k=1}^{n} I_kR_k = \sum_{i=1}^{m} \mathscr{E}_i, \tag{7.4}$$

где n — число неразветвленных участков в рассматриваемом контуре (совпадающее с числом встречающихся в этом контуре узлов), а m — число э. д. с., действующих в контуре.

Теперь можно сформулировать общие правила расчета произвольных разветвленных цепей постоянного тока.

1. Обозначить на схеме токи во всех неразветвленных участках, произвольно задавая им направление.

2. Согласно первому правилу Кирхгофа написать уравнения (7.1) для всех узлов, кроме одного (уравнение для последнего узла писать не нужно, так как оно является следствием предыдущих).

3. Согласно второму правилу Кирхгофа составить уравнения (7.4) для всех простых контуров, которые можно выделить в данной цепи и которые не получаются наложением уже рассмотренных. Простым считается такой контур, при обходе которого мы побываем в каждой точке только по одному разу. В правильно выбранной системе контуров каждый участок цепи должен фигурировать по крайней мере в одном из контуров.

4. Если в результате решения получившейся системы уравнений какие-либо токи окажутся отрицательными, то в действительности их направление противоположно выбранному на схеме.

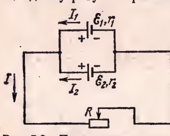


Рис. 7.2. Параллельное соединение источников тока.

Для иллюстрации применения правил Кирхгофа рассмотрим условия работы батареи из двух параллельно соединенных источников (рис. 7.2). Параметры схемы указаны на рисунке.

Обозначим токи в неразветвленных участках цепи через I_1, I_2 и I и зададим им направления, как указано на рисунке. Пользуясь сформулированными правилами расчета цепей, составляем уравнение для токов в узле:

$$I_1+I_2-I=0. \qquad (7.5)$$

Выберем два простых контура, например, следующим образом: один содержит источник с э. д. с. $\mathscr{E}_1$ и сопротивление R, другой содержит оба источника тока. Обходя эти контуры против часовой стрелки, получаем следующие уравнения:

$$I_1r_1+IR=\mathscr{E}_1, \qquad (7.6)$$

$$I_1r_1-I_2r_2=\mathscr{E}_1-\mathscr{E}_2. \qquad (7.7)$$

Мы получили систему из трех уравнений с тремя неизвестными токами I_1, I_2 и I. Выражаем I_2 из уравнения (7.5) и подставляем в (7.7):

$$I_1(r_1+r_2)-Ir_2=\mathscr{E}_1-\mathscr{E}_2 \qquad (7.8)$$

Умножая (7.6) на r_2, а (7.8) на R и складывая их почленно, находим I_1:

$$I_1=\frac{\mathscr{E}_1(r_2+R)-\mathscr{E}_2R}{R(r_1+r_2)+r_1r_2}. \qquad (7.9)$$

Исключая аналогично из уравнений (7.6) и (7.8) I_1, находим I:

$$I=\frac{\mathscr{E}_1r_2+\mathscr{E}_2r_1}{R(r_1+r_2)+r_1r_2}. \qquad (7.10)$$

Выражение для I_2, которое просто находится из уравнения (7.5), можно, учитывая симметрию схемы, написать и непосредственно, заменяя в (7.9) индексы $1\rightleftarrows 2$:

$$I_2=\frac{\mathscr{E}_2(r_1+R)-\mathscr{E}_1R}{R(r_1+r_2)+r_1r_2}. \qquad (7.11)$$

Выражение (7.10) показывает, что ток I через нагрузку R всегда положителен и, следовательно, течет в направлении, указанном на схеме. Из (7.9) и (7.11) видно, что при равных $\mathscr{E}_1$ и $\mathscr{E}_2$ токи через источники I_1 и I_2 тоже положительны. Если же, например, $\mathscr{E}_1>\mathscr{E}_2$, то I_1 всегда положителен, в то время как ток I_2 может быть и отрицательным. В таком случае он течет в направлении, противоположном указанному на рис. 7.2. Это означает, что источник $\mathscr{E}_2$ не отдает энергию во внешнюю цепь, а сам потребляет энергию от источника $\mathscr{E}_1$.

Выясним, при каких условиях источник $\mathscr{E}_2$ (при $\mathscr{E}_1>\mathscr{E}_2$) будет работать нормально, т. е. отдавать энергию во внешнюю цепь. Из (7.11) видно, что $I_2>0$ при $\mathscr{E}_2(r_1+R)-\mathscr{E}_1R>0$. Переписав это условие в виде

$$\mathscr{E}_2>\frac{\mathscr{E}_1R}{r_1+R}, \qquad (7.12)$$

видим, что второй источник работает нормально, если его э. д. с. больше напряжения U на зажимах первого источника в схеме, где только первый источник замкнут на сопротивление R. Полученный результат легко понять из следующих простых соображений. Пусть подключен к сопротивлению R только первый источник. Если напряжение на его зажимах $U=\mathscr{E}_1R/(r_1+R)$ больше $\mathscr{E}_2$, то, подключая $\mathscr{E}_2$ параллельно $\mathscr{E}_1$, мы фактически ставим второй источник на «зарядку».

Из формулы (7.12) видно, что условие нормальной работы второго источника зависит от сопротивления нагрузки R: при малом R он работает нормально, при некотором значении R, определяемом условием $\mathscr{E}_2=\mathscr{E}_1R/(r_1+R)$, ток через $\mathscr{E}_2$ обращается в нуль, т. е. при $R=\mathscr{E}_2r_1/(\mathscr{E}_1-\mathscr{E}_2)$ подключение или отключение этого источника ничего не меняет в остальной цепи. При бо́льших значениях R подключение источника $\mathscr{E}_2$ приводит к уменьшению тока через нагрузку R.

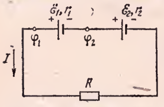


Рис. 7.3. Последовательное соединение источников тока.

Два параллельно соединенных источника тока можно заменить одним эквивалентным источником, который обеспечит во внешней цепи такой же ток. Параметры такого источника легко определить с помощью формулы (7.10). Переписывая ее в виде

$$I=\frac{\dfrac{\mathscr{E}_1r_2+\mathscr{E}_2r_1}{r_1+r_2}}{R+\dfrac{r_1r_2}{r_1+r_2}}$$

и сравнивая с выражением для тока, создаваемого эквивалентным источником э. д. с. $\mathscr{E}$ и внутренним сопротивлением, r, $I=\mathscr{E}/(R+r)$, находим, что значения $\mathscr{E}$ и r эквивалентного источника определяются формулами

$$\mathscr{E}=\frac{\mathscr{E}_1r_2+\mathscr{E}_2r_1}{r_1+r_2},\quad r=\frac{r_1r_2}{r_1+r_2}. \qquad (7.13)$$

В частности, для одинаковых параллельно соединенных источников эквивалентный источник, как видно из (7.13), имеет ту же э. д. с. и вдвое меньшее внутреннее сопротивление. В случае неравных э. д. с. $\mathscr{E}_1$ и $\mathscr{E}_2$ величина $\mathscr{E}$ имеет промежуточное значение.

Рассмотрим теперь условия работы последовательно соединенных источников тока. Для этого исследуем цепь, схема которой показана на рис. 7.3. Выясним, всегда ли наличие второго источника с э. д. с. $\mathscr{E}_2$ приводит к увеличению тока в цепи, первоначально содержащей только один источник с э. д. с. $\mathscr{E}_1$. Очевидно, что второй источник имеет

смысл подключать, только если

$$\frac{\mathscr{E}_1+\mathscr{E}_2}{r_1+r_2+R} > \frac{\mathscr{E}_1}{r_1+R}. \qquad (7.14)$$

Умножая обе части этого неравенства на положительную величину $(r_1+r_2+R)(r_1+R)$ и приводя подобные члены, получаем

$$\mathscr{E}_2(r_1+R) > \mathscr{E}_1 r_2,$$

откуда

$$\frac{\mathscr{E}_2}{r_2} > \frac{\mathscr{E}_1}{r_1+R}. \qquad (7.15)$$

Смысл этого неравенства очевиден: слева стоит ток короткого замыкания второго источника, а справа — ток в цепи, содержащей только первый источник. Итак, последовательное подключение второго источника целесообразно только в том случае, когда ток его короткого замыкания больше тока в цепи, в которую мы его собираемся включить.

§ 8. Магнитное поле постоянного тока

Движущийся электрический заряд наряду с электрическим создает магнитное поле. В отличие от потенциального электрического поля, постоянное магнитное поле, создаваемое стационарными токами, является соленоидальным, или вихревым: его силовые линии всегда замкнуты. Другими словами, магнитное поле не имеет источников — магнитных зарядов.

Магнитное поле проявляется в действии на магнитную стрелку, рамку с током, движущийся заряд. На рамку с током и магнитную стрелку магнитное поле оказывает ориентирующее действие, на движущийся заряд в магнитном поле действует сила, перпендикулярная скорости заряда.

Силовой характеристикой магнитного поля является индукция $\boldsymbol{B}$. Эта векторная физическая величина обычно вводится путем рассмотрения действия магнитного поля на маленькую пробную рамку с током. Направление вектора $\boldsymbol{B}$ совпадает с направлением нормали к свободной пробной рамке с током, установившейся в поле. За направление нормали к плоскости рамки принимают то направление, в котором будет перемещаться винт с правой нарезкой, если

вращать его по направлению тока в рамке. Величина индукции пропорциональна максимальному моменту сил, действующих на пробную рамку.

Индукция магнитного поля, создаваемого текущим по проводу током, определяется совместным действием всех отдельных участков провода. Магнитное поле удовлетворяет принципу суперпозиции: если магнитное поле создается несколькими проводниками с током, то индукция результирующего поля есть векторная сумма индукций полей, создаваемых каждым проводником. Точно так же для одного проводника с током наблюдаемая на опыте индукция $\boldsymbol{B}$ есть векторная сумма элементарных индукций $\Delta\boldsymbol{B}$, создаваемых отдельными участками провода. На опыте невозможно осуществить отдельный участок тока, так что нельзя непосредственно измерить и создаваемое им поле. Измерить можно только суммарную индукцию магнитного поля, создаваемого всеми элементами тока. Однако существует закон, называемый законом Био — Савара — Лапласа, который, будучи применен к участкам провода произвольной формы, позволяет во всех случаях вычислить значение результирующей индукции магнитного поля, совпадающее с измеренным на опыте. Закон Био — Савара — Лапласа формулируется следующим образом. Элемент провода Δl, по которому течет ток I, создает в вакууме (или в среде с магнитной проницаемостью $\mu=1$) в некоторой точке магнитное поле, индукция которого ΔB обратно пропорциональна квадрату расстояния r от элемента тока до точки наблюдения:

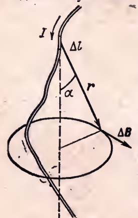


Рис. 8.1. К закону Био — Савара — Лапласа.

$$\Delta B=\frac{\mu_0}{4\pi}\frac{I\Delta l\sin\alpha}{r^2}, \tag{8.1}$$

где μ_0 — магнитная постоянная, а α — угол между направлением на точку наблюдения и направлением элемента тока Δl (рис. 8.1). Вектор $\Delta\boldsymbol{B}$ перпендикулярен к плоскости, содержащей элемент Δl и радиус-вектор $\boldsymbol{r}$. Направление

$\Delta\boldsymbol{B}$ определяется правилом правого винта: оно совпадает с направлением вращения головки винта при его поступательном перемещенни вдоль тока. Используя понятие векторного произведения, закон Био — Савара — Лапласа можно перепнсать в векторном виде:

$$\Delta\boldsymbol{B}=\frac{\mu_0 I}{4\pi}\frac{\Delta\boldsymbol{l}\times\boldsymbol{r}}{r^3}. \qquad (8.2)$$

Здесь вектор $\Delta\boldsymbol{l}$ направлен вдоль провода в направлении движения положительных зарядов.

Формула (8.1) (нлн (8.2)) позволяет рассчнтать индукцию магнитного поля, создаваемого пронзвольным распределеннем постоянных токов. Простейшнм примером нспользовання закона Био — Савара — Лапласа служнт вычнсление магнитного поля в центре кругового тока. Пусть ток I идет по проводу в виде окружностн радиуса R по часовой стрелке (рнс. 8.2). Векторы $\Delta\boldsymbol{B}$ от всех элементов кольцевого провода направлены перпендикулярно плоскостн круга за плоскость рнсунка. Поэтому суммарная нндукция магнитного поля $\boldsymbol{B}$ направлена в ту же сторону, а ее величнна равна просто сумме всех ΔB. Любой элемент кругового контура находится на одном и том же расстоянни $r=R$ от центра круга, а его направление образует прямой угол $\alpha=\pi/2$ с направлением на точку наблюдения. Поэтому, суммируя элементарные индукции, с помощью формулы (8.1) получнм

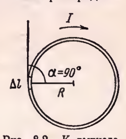


Рис. 8.2. К вычислению магнитного поля кругового тока.

$$B=\sum\Delta B=\frac{\mu_0}{4\pi}\frac{I}{R^2}\sum\Delta l.$$

Сумма длин всех элементарных участков $\sum\Delta l$ равна длине окружности $2\pi R$, поэтому нндукция магнитного поля в центре кругового тока равна

$$B=\frac{\mu_0 I}{2R}.$$

Расчет магнитного поля, создаваемого токами других конфигураций, выполняется с помощью интегрирования.

В частности, индукция магнитного поля, создаваемого бесконечным прямолинейным проводником с током, убывает обратно пропорционально расстоянию r от провода:

$$B = \frac{\mu_0}{2\pi} \frac{I}{r}. \tag{8.3}$$

Линии магнитной индукции в этом случае представляют собой концентрические окружности, плоскости которых перпендикулярны току, а центры расположены на оси тока (рис. 8.3).

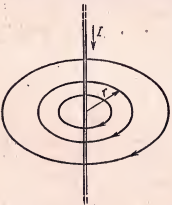


Рис. 8.3. Линии магнитной индукции прямолинейного проводника с током.

Магнитное поле может быть охарактеризовано некоторым общим соотношением, которое, как и теорема Гаусса в электростатике, может быть использовано для расчета магнитных полей, создаваемых симметричными распределениями токов. Это соотношение носит название теоремы о циркуляции вектора магнитной индукции.

Рассмотрим произвольный замкнутый контур l и зададим на нем направление обхода. Обозначим через B_l проекцию вектора $\boldsymbol{B}$ на направление элемента контура Δl. Составим сумму произведений $B_l \Delta l$ для всех элементов замкнутого контура. Эта сумма $\sum B_l \Delta l$ называется циркуляцией вектора $\boldsymbol{B}$ по замкнутому контуру l. В силу закона Био — Савара — Лапласа циркуляция вектора $\boldsymbol{B}$ по произвольному замкнутому контуру равна произведению μ_0 на ток I, пронизывающий контур, по которому берется циркуляция.

Проверим справедливость этого утверждения для магнитного поля, создаваемого прямолинейным бесконечным проводником с током. Прежде всего отметим, что нужно рассматривать только контуры, лежащие в плоскости, перпендикулярной к проводнику, так как вектор $\boldsymbol{B}$ не имеет составляющих вдоль тока и, следовательно, циркуляция $\boldsymbol{B}$ по произвольному контуру совпадает с циркуля-

цией по проекции контура на эту плоскость. Проще всего рассчитать циркуляцию $\boldsymbol{B}$ по круговому контуру с центром на проводнике. В этом случае вектор $\boldsymbol{B}$ в каждой точке контура параллелен элементу Δl (если выбранное направление обхода совпадает с направлением силовых линий), а величина B, одинаковая во всех точках контура, дается формулой (8.3). Суммируя $B_l \Delta l$ по всем элементам контура, получаем

$$\sum B_l \Delta l = B \sum \Delta l = \frac{\mu_0}{2\pi} \frac{I}{R} 2\pi R = \mu_0 I. \quad (8.4)$$

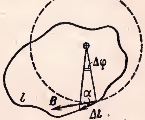


Рис. 8.4. К теореме о циркуляции вектора индукции магнитного поля.

Видно, что циркуляция $\boldsymbol{B}$ не зависит от радиуса окружности. Нетрудно убедиться в том, что при произвольной деформации окружности величина циркуляции $\boldsymbol{B}$ не изменится. Рассмотрим элемент Δl произвольного контура l (рис. 8.4). Для него $B_l \Delta l = B \Delta l \cos \alpha$; но $\Delta l \cos \alpha = r \Delta\varphi$, поэтому

$$B_l \Delta l = \frac{\mu_0 I}{2\pi r} \cdot r \Delta\varphi = \frac{\mu_0 I}{2\pi} \Delta\varphi.$$

Суммируя по всем элементам контура, получаем

$$\sum B_l \Delta l = \frac{\mu_0 I}{2\pi} \sum \Delta\varphi = \mu_0 I. \quad (8.5)$$

Теорема о циркуляции вектора индукции магнитного поля справедлива для поля, создаваемого произвольным распределением токов.

Применим теорему о циркуляции вектора индукции магнитного поля к расчету поля, создаваемого соленоидом, т. е. цилиндрической катушкой с плотно соприкасающимися витками. Магнитное поле такой катушки имеет вид, показанный на рис. 8.5. Если длина катушки много больше ее диаметра, то линии магнитной индукции внутри катушки параллельны ее оси и поле там однородно всюду, за исключением концов катушки. Снаружи вблизи боковой поверхности катушки поле практически отсутствует. Вычислим циркуляцию индукции $\boldsymbol{B}$ по прямоугольному

контуру, показанному на рис. 8.6: сторона bc параллельна, а стороны ab и cd перпендикулярны линиям индукции внутри катушки. Тогда вектор $\boldsymbol{B}$ будет иметь отличную от нуля проекцию на направление контура только на участке bc и циркуляция $\boldsymbol{B}$ по контуру равна Bl, где l — длина

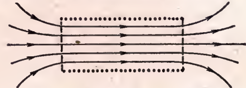


Рис. 8.5. Магнитное поле соленоида.

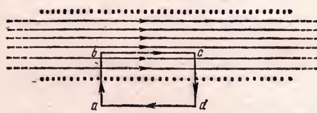


Рис. 8.6. Контур для применения теоремы о циркуляции.

участка bc. Подсчитаем теперь полный ток, пронизывающий выбранный контур. Обозначим число витков на единицу длины соленоида через n. Сквозь выбранный контур проходит nl витков, и полный ток равен Inl. Согласно теореме о циркуляции

$$Bl = \mu_0 Inl,$$

откуда

$$B = \mu_0 In. \tag{8.6}$$

Формула (8.6) дает индукцию магнитного поля внутри длинного соленоида, по обмотке которого пропускается ток I. Вблизи краев соленоида поле уже не будет однородным. Легко показать, что индукция поля на оси соленоида на самом его конце равна половине значения индукции внутри соленоида. Если к концу соленоида приставить другой такой же соленоид, по которому в том же направлении протекает такой же ток, то рассматриваемая точка окажется

внутри нового, составного соленоида, и индукция поля в ней будет определяться формулой (8.6). Но по принципу суперпозиции эта же индукция есть сумма индукций полей, существующих вблизи концов каждого соленоида. Поскольку соленоиды одинаковы, то одинаковы и создаваемые ими поля, и, следовательно, индукция магнитного поля в точке на оси на конце одного соленоида равна $B = {}^{1}/_{2}\mu_0 In$.

Теперь вычислим индукцию магнитного поля внутри замкнутой тороидальной катушки (рис. 8.7). В отличие от соленоида, линии магнитной индукции замыкаются здесь внутри самой катушки и представляют собой окружности, параллельные оси тора. Направление их таково, что, глядя вдоль них, мы видим токи в обмотке тороидальной катушки текущими по часовой стрелке. Вычислим циркуляцию вектора магнитной индукции вдоль одной из таких линий. Из соображений симметрии очевидно, что величина вектора индукции $\boldsymbol{B}$ одинакова во всех точках, лежащих на одной линии индукции. Пусть радиус такой окружности равен r $(r_1 < r < r_2)$. Тогда по теореме о циркуляции имеем

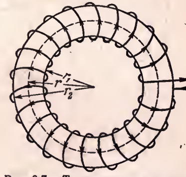


Рис. 8.7. Тороидальная катушка.

$$B \cdot 2\pi r = \mu_0 IN,$$

где N — полное число витков, а I — ток в обмотке тороидальной катушки. Отсюда

$$B = \frac{\mu_0}{2\pi}\frac{IN}{r}. \qquad (8.7)$$

Формула (8.7) показывает, что индукция магнитного поля в торе максимальна вблизи внутренней стороны и минимальна вблизи внешней стороны тора.

Выражение для индукции магнитного поля в длинном соленоиде (8.6) может быть получено как предельный случай формулы (8.7) для поля в тороидальной катушке при условии, что диаметр витков много меньше радиуса самого тора. В этом случае $(r_2 - r_1)/r \ll 1$, поле внутри тора

практически однородно, а отношение $N/2\pi r$ представляет собой число витков на единицу длины катушки.

В заключение этого параграфа вернемся к формуле (8.3) для индукции магнитного поля прямолинейного

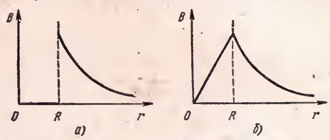


Рис. 8.8. Магнитное поле прямого тока, текущего по поверхности цилиндрического проводника (*а*) и по всему сечению (*б*).

бесконечного проводника с током. Для очень тонкого проводника, когда r стремится к нулю, индукция магнитного поля вблизи проводника неограниченно возрастает. Реально провод всегда имеет конечную толщину. С помощью теоремы о циркуляции индукции магнитного поля легко убедиться, что снаружи проводника индукция поля по-прежнему выражается формулой (8.3), а внутри проводника величина индукции зависит от распределения тока по сечению проводника. Если весь ток течет только по поверхности цилиндрического проводника, как это бывает в полой тонкостенной трубке или в сверхпроводниках, то магнитного поля внутри нет. Зависимость величины индукции от расстояния r до оси проводника имеет в этом случае вид, показанный на рис. 8.8, *а*. Если ток равномерно распределен по сечению проводника, то магнитное поле внутри проводника пропорционально расстоянию от его оси (рис. 8.8, *б*). Чтобы убедиться в этом, рассмотрим циркуляцию $\boldsymbol{B}$ по кру-

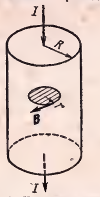


Рис. 8.9. К расчету магнитного поля внутри проводника.

говому контуру радиуса r, лежащему внутри проводника в плоскости, перпендикулярной к его оси (рис. 8.9). По соображениям симметрии величина индукции поля одинакова для всех точек, лежащих на такой окружности. Так как поле направлено по касательной к окружности, то по теореме о циркуляции имеем

$$B \cdot 2\pi r = \mu_0 I', \tag{8.8}$$

где I' есть ток, проходящий через заштрихованное на рисунке сечение внутри этой окружности. При равномерном распределении тока по сечению

$$I' = I\left(\frac{r}{R}\right)^2,$$

поэтому согласно (8.8) индукция поля внутри проводника равна

$$B = \frac{\mu_0}{2\pi} I \frac{r}{R^2}. \tag{8.9}$$

ЭЛЕКТРОМАГНИТНОЕ ПОЛЕ

§ 9. Явление электромагнитной индукции. Самоиндукция. Энергия магнитного поля

Открытое в 1831 году Фарадеем явление электромагнитной индукции состоит в том, что в любом замкнутом контуре при изменении магнитного потока через поверхность, ограниченную этим контуром, возникает электродвижущая сила, величина которой пропорциональна скорости изменения магнитного потока:

$$\mathscr{E} = -k\frac{d\Phi}{dt}. \tag{9.1}$$

Коэффициент k зависит от выбора системы единиц. В единицах СИ $k=1$.

В проводящем контуре существование э. д. с. индукции приводит к появлению индукционного тока. Знак минус в формуле (9.1) соответствует правилу Ленца, согласно которому направление индукционного тока всегда таково, что создаваемое им магнитное поле препятствует изменению магнитного потока, вызывающего индукционный ток.

Рассматривая различные опыты, можно убедиться, что возникновение электродвижущей силы индукции может

быть обусловлено сторонними силами разной физической природы. В неподвижном контуре э. д. с. индукции обусловлена вихревым электрическим полем, возникающим при изменении магнитного потока через поверхность, ограниченную контуром. Изменение магнитного потока может быть вызвано как движением магнита, создающего поле, так и изменением величины магнитного поля путем изменения тока в электромагните. Первая возможность реализуется в промышленных генераторах, где вращающийся электромагнит возбуждает ток в обмотках неподвижного статора. Вторая возможность реализуется в трансформаторах, где изменение тока в первичной обмотке вызывает изменение магнитного потока и, следовательно, появление вихревого электрического поля.

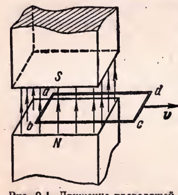


Рис. 9.1. Движение проводящей рамки в магнитном поле.

В отличие от потенциального электростатического поля, создаваемого неподвижными электрическими зарядами, вихревое электрическое поле, возникающее при изменении магнитного поля, обладает тем свойством, что работа сил этого поля на замкнутом пути не равна нулю. Именно этой работой и определяется э. д. с. индукции в замкнутом контуре. Подчеркнем, что вихревое электрическое поле при изменении магнитного поля существует независимо от того, имеется ли в этом месте замкнутый проводящий контур. Сам проводящий контур является лишь индикатором, обнаруживающим наличие этого вихревого электрического поля.

Иную физическую природу имеет вызывающая индукционный ток сторонняя сила, возникающая при движении проводника в неизменном магнитном поле. Вихревое электрическое поле в этом случае отсутствует, а сторонняя сила есть сила Лоренца, с которой магнитное поле действует на движущиеся вместе с проводником электрические заряды. На таком принципе основано действие электрических генераторов небольшой мощности, где индукционный ток воз-

буждается в обмотке ротора, вращающегося в неподвижном магнитном поле. В отсутствие проводника, содержащего электрические заряды, никаких сторонних сил, а следовательно и э. д. с. индукции, нет.

Нетрудно убедиться, что э. д. с. индукции, вычисляемая по общему закону (9.1), совпадает с работой сторонней силы (силы Лоренца) по перемещению единичного заряда по замкнутому контуру. Будем считать, что прямоугольная металлическая рамка *abcd* движется с постоянной скоростью v, как показано на рис. 9.1. Сторона *ab* пересекает силовые линии однородного магнитного поля B, существующего между полюсами магнита. Величина э. д. с. индукции в контуре *abcd* может быть вычислена с помощью закона электромагнитной индукции (9.1). Учитывая, что при движении рамки в направлении, указанном на рис. 9.1, пронизывающий рамку магнитный поток убывает, имеем

$$\Delta\Phi = -Blv\,\Delta t,$$

откуда согласно (9.1)

$$\mathscr{E} = -\frac{d\Phi}{dt} = Blv. \qquad (9.2)$$

Вычислим теперь эту же э. д. с. индукции как работу сторонних сил при перемещении единичного заряда по контуру *abcd*. Сторонняя сила — сила Лоренца — дает вклад в э. д. с. только на участке *ab* длиной l (рис. 9.1). Отношение силы к переносимому заряду равно vB. Поэтому величина э. д. с. индукции равна Blv, что совпадает с (9.2). Отметим, что э. д. с. индукции в такой рамке возникает лишь тогда, когда только часть рамки находится в однородном магнитном поле (как показано на рисунке) или когда магнитное поле неоднородно. Если рамка целиком находится в однородном поле, то при ее поступательном движении магнитный поток не меняется и э. д. с. индукции равна нулю, хотя на отдельных участках рамки сторонние силы действуют.

Несмотря на разную физическую природу сторонних сил в рассмотренных случаях, закон электромагнитной индукции (9.1), согласно которому э. д. с. индукции в контуре равна взятой с обратным знаком скорости изменения магнитного потока, справедлив и в случае, когда поток меняется за счет изменения магнитного поля, и в случае, когда

поток меняется за счет движения контура в неизменном магнитном поле, и в случае, когда происходит и то и другое. Эти возможности — поле меняется или контур движется — неразличимы в формулировке закона индукции. Рассмотрим эти возможности с точки зрения сторонних сил, действующих на заряды.

Сила, действующая на заряд q в электрическом поле напряженности $\boldsymbol{E}$, равна $q\boldsymbol{E}$ независимо от того, является поле потенциальным или вихревым, т. е. создается электрическими зарядами или изменяющимся магнитным полем.

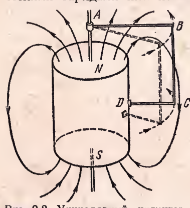


Рис. 9.2. Униполярный индуктор.

Сила Лоренца, действующая на движущийся со скоростью $\boldsymbol{v}$ заряд q в магнитном поле индукции $\boldsymbol{B}$, определяется векторным произведением скорости $\boldsymbol{v}$ и индукции $\boldsymbol{B}$:

$$\boldsymbol{F} = q\boldsymbol{v} \times \boldsymbol{B}. \quad (9.3)$$

Эта сила перпендикулярна как скорости заряда, так и индукции магнитного поля. Полная сила, действующая на заряд q, равна

$$\boldsymbol{F} = q(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}). \quad (9.4)$$

В движущихся в магнитном поле проводниках сила возникает за счет $\boldsymbol{v}$. Вихревое электрическое поле возникает, если где-либо меняется магнитное поле. Эти эффекты независимы и могут проявляться и порознь, и одновременно. Но в любом случае действие этих сторонних сил создает в контуре электродвижущую силу, величина которой равна скорости изменения магнитного потока.

В этом смысле формулу (9.1) можно назвать правилом для нахождения э. д. с. индукции, ибо, как мы видели, эта формула не вскрывает физической причины возникновения э. д. с.— причина может быть разной в разных случаях. Правило потока (9.1) дает только значение э. д. с., механизм ее возникновения должен устанавливаться независимо.

Но из этого правила бывают и исключения! Понять эти исключения можно, лишь зная истинную причину возникновения сторонних сил. Яркий пример — известный еще

Фарадею униполярный индуктор (рис. 9.2). Металлический контур $ABCD$ вращается вокруг постоянного магнита цилиндрической формы, образуя с магнитом замкнутую электрическую цепь при помощи двух скользящих контактов, один из которых (A) касается оси магнита, а другой (D) — самого магнита в нейтральной точке. Магнитный поток через контур $ABCD$, замыкаемый частью магнита, равен нулю в любой момент времени, так как силовые

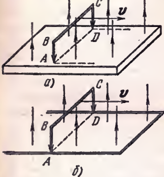


Рис. 9.3. К объяснению действия униполярного индуктора.

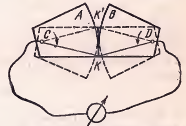


Рис. 9.4. При повороте пластин вокруг точек C и D точка контакта перемещается из K в K'.

линии магнитного поля лежат в плоскости контура. Изменение магнитного потока при вращении контура также равно нулю, а индукционный ток есть! Если отчетливо представить себе, что причиной возникновения э. д. с. в таком устройстве является сила Лоренца, действующая на электроны в движущемся контуре, то легко сообразить, что э. д. с. действительно должна возникать. Для большей наглядности рассмотрим видоизмененный вариант этого опыта, когда контур $ABCD$ движется поступательно вдоль проводящей ленты, помещенной в однородное магнитное поле (рис. 9.3, *а*). Здесь также магнитный поток через контур и его изменение равны нулю, но на участке BC на электроны действует сила Лоренца, заставляющая их двигаться вдоль контура. Возникающая э. д. с. $\mathscr{E} = Blv$ будет такой же, как и в устройстве, показанном на рис. 9.3, *б*, где вместо ленты имеются проводящие рельсы, соединенные в одном месте.

А вот в устройстве, показанном на рис. 9.4 (магнитное поле направлено перпендикулярно плоскости рисунка), при небольшом повороте пластин A и B вокруг точек C и D

точка контакта K перемещается на большое расстояние, площадь контура и, следовательно, магнитный поток изменяются на значительную величину, но тока в цепи практически нет! И легко понять, почему. Перемещение точки контакта происходит от очень незначительного движения пластин, тогда $\boldsymbol{v}\times\boldsymbol{B}$ очень мало и сила Лоренца практически отсутствует.

Важным частным случаем электромагнитной индукции является самоиндукция, когда изменяющийся магнитный поток, вызывающий э. д. с. индукции, создается током в самом рассматриваемом контуре. Согласно правилу Ленца явление самоиндукции препятствует изменению тока в контуре. Поэтому при замыкании цепи, содержащей источник постоянного тока, ток достигает своего установившегося значения не сразу, а при размыкании цепи не может мгновенно исчезнуть.

Рассмотрим, от чего зависит э. д. с. самоиндукции. Магнитное поле, создаваемое током в контуре (или катушке), в любой точке пропорционально силе тока I. Поэтому и магнитный поток Φ, пронизывающий этот контур, пропорционален току:

$$\Phi = LI. \tag{9.5}$$

Коэффициент пропорциональности L называется индуктивностью контура или коэффициентом самоиндукции. Используя закон электромагнитной индукции (9.1), для э. д. с. самоиндукции получаем выражение

$$\mathscr{E} = -L\frac{dI}{dt}. \tag{9.6}$$

При неизменных форме и размерах контура э. д. с. самоиндукции пропорциональна скорости изменения тока в контуре. Индуктивность контура L зависит от его размеров и формы, а также от магнитных свойств среды, в которую он помещен.

Для примера найдем индуктивность длинного соленоида, имеющего N витков, площадь сечения S и длину l. Индукция магнитного поля внутри такого соленоида согласно (8.6) равна

$$B = \mu_0 \frac{N}{l} I. \tag{9.7}$$

Магнитный поток Φ_1 через поперечное сечение соленоида S равен

$$\Phi_1 = BS = \mu_0 \frac{N}{l} SI,$$

а суммарный поток Φ через все N витков будет в N раз больше:

$$\Phi = N\Phi_1 = \mu_0 \frac{N^2}{l} SI. \qquad (9.8)$$

Сравнивая (9.8) и (9.5), получаем

$$L = \mu_0 \frac{N^2 S}{l}. \qquad (9.9)$$

Вводя число витков на единицу длины $n = N/l$, этому выражению можно придать вид

$$L = \mu_0 n^2 V,$$

где $V = Sl$ — объем соленоида, внутри которого в основном и сосредоточено магнитное поле.

Магнитное поле обладает энергией. Проще всего в этом убедиться, рассматривая процесс спадания тока в катушке

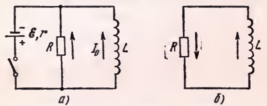


Рис. 9.5. Электрическая цепь для изучения магнитной энергии тока.

при отсоединении ее от источника тока в схеме на рис. 9.5, *а*. До размыкания ключа в катушке идет некоторый ток I_0, и этот ток создает магнитное поле. При размыкании ключа остается последовательная цепь из катушки и сопротивления R (рис. 9.5, *б*). Ток в катушке благодаря самоиндукции спадает постепенно, и при этом на сопротивлении R продолжает выделяться джоулево тепло. За счет каких запасов энергии выделяется тепло — ведь источник питания уже отключен? Здесь убывает ток и создаваемое им магнитное поле; значит, мы можем говорить об энергии тока или об

энергии создаваемого им магнитного поля. По аналогии с электростатикой, где можно говорить об энергии зарядов или об энергии создаваемого ими поля, естественно ожидать, что в случае постоянного тока допустимы оба представления: энергию можно рассматривать либо как энергию тока, либо как энергию создаваемого им магнитного поля. Но мы уже видели, что, хотя не бывает электрического заряда без создаваемого им поля, электрическое поле без заряда — вихревое поле — может существовать и оно обладает энергией. Поэтому вопрос о локализации электрической энергии решается в пользу поля. Как мы увидим немного позже, точно так же обстоит дело и с магнитной энергией.

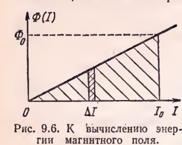


Рис. 9.6. К вычислению энергии магнитного поля.

Подсчитаем энергию магнитного поля. Из закона сохранения энергии очевидно, что в рассматриваемом нами опыте (рис. 9.5, *б*) вся энергия магнитного поля в конце концов выделится в виде джоулева тепла на сопротивлении R. За время Δt на сопротивлении R выделяется тепло $\Delta Q = I^2 R \Delta t$. По закону Ома ток I равен

$$I = \frac{\mathscr{E}}{R} = -\frac{L}{R}\frac{\Delta I}{\Delta t}.$$

С учетом этого равенства выражение для ΔQ можно записать в виде

$$\Delta Q = -LI\,\Delta I = -\Phi(I)\,\Delta I. \qquad (9.10)$$

Выделяющееся тепло ΔQ, разумеется, положительно, так как ток убывает и, следовательно, $\Delta I < 0$. Изобразив на графике зависимость магнитного потока $\Phi(I)$ от тока I (рис. 9.6), легко сообразить, что полное выделившееся тепло, равное первоначальному запасу энергии магнитного поля, определяется площадью заштрихованного треугольника $^1/_2\Phi_0 I_0$. Таким образом, выражение для энергии магнитного поля W, создаваемого током I в катушке с индуктивностью L, имеет вид

$$W = \frac{1}{2}\Phi I = \frac{1}{2}LI^2 = \frac{\Phi^2}{2L}. \qquad (9.11)$$

Как и в электростатике, можно ввести понятие объемной плотности энергии магнитного поля. Рассматривая однородное магнитное поле внутри длинного соленоида, подставим во вторую из формул (9.11) выражение для индуктивности (9.9), а ток I выразим через индукцию магнитного поля с помощью (9.7). В результате получим

$$W = \frac{1}{2\mu_0} B^2 V,$$

откуда объемная плотность энергии магнитного поля w равна

$$w = \frac{B^2}{2\mu_0}. \qquad (9.12)$$

§ 10. Относительный характер электрического и магнитного полей. Основы теории электромагнитного поля

Изучая электромагнитную индукцию, мы видели, что при рассмотрении этого явления в определенной инерциальной системе отсчета возможны две различные причины возникновения индукционного тока: либо появление вихревого электрического поля, либо действие силы Лоренца на движущиеся вместе с проводником электрические заряды со стороны магнитного поля. Однако при анализе возникновения э. д. с. индукции за счет силы Лоренца в опыте с металлической рамкой, движущейся в магнитном поле (рис. 9.1), мы можем рассуждать и иначе.

Перейдем в систему отсчета, связанную с движущейся рамкой. В этой системе отсчета заряды неподвижны и, следовательно, со стороны магнитного поля сила на них не действует.

Как же объяснить возникновение э. д. с. индукции в этой системе отсчета? Единственное, что остается предположить,— это наличие в этой системе электрического поля, направленного перпендикулярно магнитному вдоль стороны ab рамки, которого не было в исходной системе отсчета. Действительно, в любой инерциальной системе отсчета действующая на заряд сила определяется формулой (9.4), и, поскольку в системе отсчета, связанной с рамкой, $\boldsymbol{v}=0$, сила $\boldsymbol{F}$ может быть обусловлена только электрическим полем $\boldsymbol{E}'$, существующим в этой системе.

Итак, мы приходим к выводу об относительном характере электрического и магнитного полей. Согласно принципу относительности, с которым подробнее познакомимся ниже (см. стр. 476), все инерциальные системы отсчета равноправны. В обсуждаемом здесь опыте наблюдаемой величиной является э. д. с. индукции в рамке, и она существует независимо

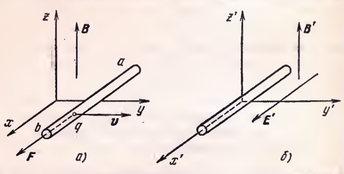


Рис. 10.1. К объяснению возникновения э. д. с. индукции в разных системах отсчета.

от того, в какой инерциальной системе этот опыт рассматривается. Как мы видели, в одной системе отсчета, где электрическое поле отсутствует, существование э. д. с. объясняется силой Лоренца (рис. 10.1, *а*), в то время как в другой, где рамка неподвижна,— только наличием электрического поля (рис. 10.1, *б*). При малых скоростях ($v \ll c$), когда можно пренебречь изменением силы $\boldsymbol{F}$ при переходе от одной системы отсчета к другой, из формулы (9.4) следует, что напряженность электрического поля $\boldsymbol{E}'$ в системе, где рамка неподвижна, должна быть равна

$$\boldsymbol{E}' = \boldsymbol{v} \times \boldsymbol{B}. \qquad (10.1)$$

Итак, движущийся магнит кроме магнитного создает и электрическое поле.

Обратим внимание на то, что относительный характер электрического и магнитного полей мы могли заметить и раньше. В самом деле, неподвижный заряд создает только электрическое поле. Однако заряд, неподвижный в какой-либо одной системе отсчета, относительно других систем отсчета движется. Такой движущийся заряд подобен

электрическому току и потому создает магнитное поле. Таким образом, если в какой-либо системе отсчета есть только электрическое поле, то в любой другой системе будет еще и магнитное. Получим формулу для индукции магнитного поля в этом случае, аналогичную формуле (10.1). Рассмотрим систему отсчета, движущуюся со скоростью $\boldsymbol{v}$ относительно заряда q. В этой системе отсчета заряд движется со скоростью $-\boldsymbol{v}$. Воспользуемся законом Био — Савара — Лапласа (8.2) для нахождения индукции магнитного поля $\boldsymbol{B}'$, создаваемого движущимся со скоростью $-\boldsymbol{v}$ зарядом q. Ток I выражается через концентрацию n элементарных зарядов q_0, скорость их движения v, площадь сечения проводника S следующим образом:

$$I = q_0 n S v.$$

Подставляя это выражение в закон Био — Савара — Лапласа (8.2) и учитывая, что векторы $\Delta \boldsymbol{l}$ и $-\boldsymbol{v}$ параллельны, получаем

$$\boldsymbol{B}' = -\frac{\mu_0}{4\pi} q_0 n S \, \Delta l \, \frac{\boldsymbol{v} \times \boldsymbol{r}}{r^3},$$

Так как $q_0 n S \Delta l$ есть полный заряд q, находящийся в рассматриваемом объеме проводника $S\Delta l$, то создаваемое этим зарядом при движении магнитное поле есть

$$\boldsymbol{B}' = -\frac{\mu_0}{4\pi} q \frac{\boldsymbol{v} \times \boldsymbol{r}}{r^3}. \qquad (10.2)$$

Но в этой же точке заряд q создает электрическое поле $\boldsymbol{E}$, равное

$$\boldsymbol{E} = \frac{q}{4\pi\varepsilon_0} \frac{\boldsymbol{r}}{r^3}. \qquad (10.3)$$

Сравнивая (10.2) и (10.3), видим, что магнитное поле, создаваемое движущимся со скоростью $-\boldsymbol{v}$ зарядом, связано с электрическим полем $\boldsymbol{E}$, создаваемым этим же зарядом в той системе отсчета, где он неподвижен, соотношением

$$\boldsymbol{B}' = -\varepsilon_0 \mu_0 \boldsymbol{v} \times \boldsymbol{E}. \qquad (10.4)$$

Эта формула, полученная для точечного заряда, справедлива и для поля, создаваемого любым распределением зарядов. Таким образом, если в некоторой системе отсчета сущест-

вует только электрическое поле $\boldsymbol{E}$, то в другой системе отсчета, движущейся со скоростью $\boldsymbol{v}$ относительно исходной, существует еще и магнитное поле $\boldsymbol{B}'$, которое вычисляется по формуле (10.4).

Формулы (10.1) и (10.4) представляют собой частные случаи преобразования полей при переходе от одной инерциальной системы отсчета к другой. Они справедливы при малой относительной скорости систем отсчета ($v \ll c$). В общем случае, когда в исходной системе отсчета есть и электрическое, и магнитное поле, нерелятивистские формулы преобразования имеют вид

$$\boldsymbol{E}' = \boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}, \quad \boldsymbol{B}' = \boldsymbol{B} - \varepsilon_0\mu_0 \boldsymbol{v} \times \boldsymbol{E}. \tag{10.5}$$

Отметим, что $\varepsilon_0\mu_0 = 1/c^2$, где c — скорость света в вакууме.

Формулы преобразования электрического и магнитного полей при относительной скорости систем отсчета, сравнимой со скоростью света, более громоздки, чем (10.5). Однако всегда при переходе от одной инерциальной системы отсчета к другой существуют инвариантные, т. е. не меняющие своего значения, комбинации из векторов $\boldsymbol{E}$ и $\boldsymbol{B}$, причем их только две—это скалярное произведение этих векторов и их разность квадратов:

$$\boldsymbol{E} \cdot \boldsymbol{B} = \boldsymbol{E}' \cdot \boldsymbol{B}', \tag{10.6}$$

$$E^2 - c^2B^2 = E'^2 - c^2B'^2. \tag{10.7}$$

Формулы (10.6), (10.7) позволяют сделать ряд важных выводов о свойствах электромагнитного поля. Если в какой-либо инерциальной системе отсчета электрическое и магнитное поля взаимно перпендикулярны, то, как видно из (10.6), они будут взаимно перпендикулярны и во всякой другой системе. Для таких взаимно ортогональных полей можно найти такую систему отсчета, в которой либо $\boldsymbol{B}=0$, либо $\boldsymbol{E}=0$, смотря по тому, положителен или отрицателен инвариант (10.7).

Из относительного характера электрического и магнитного полей естественно вытекает, что при изучении электрических и магнитных явлений имеет смысл рассматривать эти поля совместно, как единое электромагнитное поле. При переходе от одной системы отсчета к другой электрическое поле в одной системе, как мы видели, выражается и через электрическое поле, и через магнитное поле в другой системе, и наоборот. Поэтому естественно ожидать,

что между электрическими и магнитными явлениями существует определенная симметрия. Изменение магнитного поля порождает вихревое электрическое поле. Оказывается, что справедливо и обратное: изменяющееся во времени электрическое поле порождает магнитное поле.

К этому выводу можно прийти, анализируя уже известные нам экспериментальные факты и описывающие их физические законы. Рассмотрим участок электрической цепи, содержащий длинный прямолинейный провод и плоский конденсатор (рис. 10.2). Будем считать, что в течение некоторого достаточно малого промежутка времени ток в этой цепи равен I. Этот ток связан с изменением заряда конденсатора соотношением $I=dq/dt$.

Рассмотрим круговой контур l, охватывающий проводник, как показано на рис. 10.2. Ток I создает магнитное поле, поэтому по теореме о циркуляции вектора индукции магнитного поля (8.5) имеем

$$\sum_l B_l \Delta l = \mu_0 I. \qquad (10.8)$$

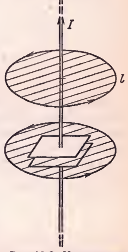


Рис. 10.2. Циркуляция вектора магнитной индукции не зависит от положения контура.

В правой части (10.8) стоит заряд, пересекающий ограниченную контуром l поверхность в единицу времени. Будем теперь перемещать контур l вниз до тех пор, пока его плоскость не пройдет в промежутке между пластинами. В этом случае никакие заряды не пересекают ограниченную контуром часть плоскости, и в этом смысле ток I в (10.8) равен нулю. Но магнитное поле вокруг провода, в том месте, где расположен контур, исчезнуть не может, и левая часть (10.8) не изменяет своего значения при смещении контура. Мы приходим к противоречию: левая часть (10.8) отлична от нуля, а правая равна нулю. Значит, в формуле (10.8) чего-то не хватает. Естественно ожидать, что на самом деле в правой части этой формулы должен стоять еще один член, который равен нулю, пока плоскость контура пересекает провод. Как угадать вид этого члена? Так как левая

часть (10.8) прн смещении контура не изменилась, то попробуем подставить в правую часть (10.8) вместо I равную ему скорость изменения заряда на обкладках конденсатора dq/dt и попытаемся интерпретировать эту величину так, чтобы она имела смысл и в той области, где отсутствуют движущиеся заряды. Поскольку заряд конденсатора q равен произведению поверхностной плотности заряда σ на площадь пластины S, то при неизменных размерах и форме конденсатора $dq/dt=S\,d\sigma/dt$. Выражая поверхностную плотность заряда через напряженность электрического поля между пластинами $E=\sigma/\varepsilon_0$, перепишем (10.8) в виде

$$\sum_l B_l\,\Delta l=\varepsilon_0\mu_0 S\frac{dE}{dt}. \qquad (10.9)$$

В отличие от тока I, величина dE/dt не равна нулю в промежутке между обкладками конденсатора. Поскольку произведение ES представляет собой поток напряженности электрического поля E через поверхность, ограниченную контуром, то в правой части (10.9) стоит величина, пропорциональная скорости изменения потока N напряженности электрического поля:

$$\sum_l B_l\,\Delta l=\varepsilon_0\mu_0\frac{dN}{dt}. \qquad (10.10)$$

Если теперь вместо (10.8) и (10.10) написать формулу

$$\sum_l B_l\,\Delta l=\mu_0 I+\varepsilon_0\mu_0\frac{dN}{dt}, \qquad (10.11)$$

то она будет справедлива всегда, независимо от того, где проходит контур l. Если плоскость контура пересекает провод, то второй член в правой части (10.11) практически равен нулю, и мы возвращаемся к теореме о циркуляции магнитного поля (10.8). Если же плоскость контура проходит внутри конденсатора, то $I=0$, первый член в правой части вклада не дает, но, как мы видели, положение спасает второе слагаемое.

Возникает вопрос: является ли добавленное нами второе слагаемое в правой части (10.11) чисто формальным, необходимым только для того, чтобы формула была спра-

ведлива для любого контура, или оно имеет физический смысл и соответствует тому, что магнитное поле возбуждается изменяющимся электрическим полем? Ответ на этот вопрос можно получить, если рассмотреть несколько видоизмененный опыт (рис. 10.3), где контур l расположен целиком внутри большого конденсатора, расстояние между пластинами которого велико по сравнению с размерами контура. Опыт показывает, что внутри конденсатора есть магнитное поле, однако очевидно, что это поле не может создаваться далеко расположенными проводами с током I. Значит, в этом случае магнитное поле возникает из-за изменения электрического поля. Циркуляция индукции этого магнитного поля по контуру определяется скоростью изменения потока напряженности электрического поля через поверхность, ограниченную этим контуром.

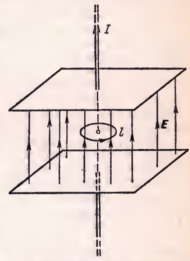


Рис. 10.3. Изменение электрического поля приводит к появлению магнитного поля.

Величина $\varepsilon_0 dN/dt$ получила название тока смещения, так как она, подобно току проводимости I, является источником магнитного поля. Термин «смещение» обусловлен историческими причинами и связан с утратившей значение механической моделью электрического поля. Следует отметить, что ток смещения эквивалентен току проводимости только в отношении способности создавать магнитное поле. Например, при наличии тока смещения не выделяется джоулево тепло.

Ток смещения был впервые предсказан Максвеллом на основе теоретического анализа известных к тому времени экспериментально установленных законов электромагнетизма. Максвелл показал, что единая непротиворечивая картина электромагнитных явлений может быть создана,

только если предположить, что изменяющееся электрическое поле способно создавать магнитное поле. Из написанной им системы уравнений электромагнитного поля следуют как все экспериментальные законы электромагнетизма, так и существование тока смещения.

Система уравнений Максвелла содержит четыре основных закона электромагнетизма. Первый закон — теорема Гаусса, связывающая поток напряженности электрического поля через замкнутую поверхность с полным зарядом внутри этой поверхности. Теорема Гаусса дает иную математическую формулировку экспериментальному закону Кулона. Второй закон — теорема Гаусса для магнитного поля, согласно которой поток вектора магнитной индукции через любую замкнутую поверхность равен нулю. Эта теорема отражает вихревой характер магнитного поля и отсутствие в природе магнитных зарядов.

Третий закон — закон электромагнитной индукции Фарадея, согласно которому изменяющееся магнитное поле порождает вихревое электрическое поле. Четвертый закон является обобщением закона Био — Савара — Лапласа. Магнитное поле может создаваться как движущимися электрическими зарядами, т. е. токами проводимости, так и изменяющимся электрическим полем, т. е. токами смещения.

Анализируя систему уравнений электромагнитного поля, Максвелл пришел к выводу, что эти уравнения допускают существование связанных между собой электрического и магнитного полей, распространяющихся в пространстве со скоростью света,— электромагнитных волн, которые позднее были экспериментально обнаружены Герцем. Разговор о свойствах этих волн нам предстоит в разделе «Колебания и волны».

§ 11. Электрические машины постоянного тока

Интересно обсудить явление электромагнитной индукции с точки зрения энергетических превращений. Для этого рассмотрим проводящий контур с включенным в него источником тока с э. д. с. $\mathcal{E}$, помещенный в однородное магнитное поле B (рис. 11.1). Подвижная часть контура AC может без трения скользить по рельсам. Обозначим через I силу тока в контуре. Тогда на подвижную часть AC действует сила Ампера $F_A = BIl$, направленная, как указано на ри-

сунке. Сила Ампера не зависит от того, покоится стержень AC или движется. Если мы хотим, чтобы стержень двигался равномерно, необходимо приложить к нему внешнюю силу F, которая в любой момент уравновешивала бы силу Ампера. Пусть за промежуток времени Δt стержень переместился на расстояние Δx в направлении действия силы Ампера. Напишем уравнение баланса энергии. За время Δt

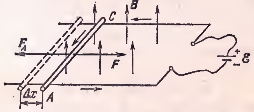


Рис. 11.1. К энергетическим превращениям при электромагнитной индукции.

источник совершил работу $\mathscr{E} I \Delta t$. При этом во всех проводниках выделилось джоулево тепло $I^2 R \Delta t$, где R — полное сопротивление цепи. Кроме того, совершила работу сила Ампера:

$$\Delta A = F_A \Delta x = IBl \Delta x = IB \Delta S = I \Delta\Phi. \qquad (11.1)$$

На основании закона сохранения энергии совершенная источником работа равна сумме выделившегося тепла и работы силы Ампера:

$$\mathscr{E} I \Delta t = I^2 R \Delta t + I \Delta\Phi, \qquad (11.2)$$

откуда, переходя к пределу $\Delta t \to 0$, для силы тока I получаем

$$I = \frac{\mathscr{E} - d\Phi/dt}{R}. \qquad (11.3)$$

Сопоставляя это выражение с законом Ома для полной цепи, мы видим, что роль э. д. с. играет величина, состоящая из двух слагаемых: из э. д. с. источника тока $\mathscr{E}$ и из величины $-d\Phi/dt$. Этот член представляет собой добавочную э. д. с. (э. д. с. индукции), обусловленную действием сторонних сил при движении участка AC.

Рассмотренный пример представляет собой модель электродвигателя постоянного тока. Как видно из уравнения

(11.2), энергия источника тока используется для совершения работы над внешними телами и частично рассеивается в виде джоулева тепла. Из формулы (11.3) видно, какую роль играет явление электромагнитной индукции в работе электромотора.

Теперь предположим, что стержень AC скользит в противоположном направлении, т. е. в направлении действия внешней силы. При этом работа силы Ампера $\Delta A = I\Delta\Phi$ отрицательна, так как $\Delta\Phi < 0$. Равная ей по величине работа внешней силы F положительна. Джоулево тепло теперь выделяется как за счет работы источника тока, так и за счет работы внешней силы F. Закон сохранения энергии имеет вид

$$\mathcal{E} I\,\Delta t + I\,|\Delta\Phi| = I^2 R\,\Delta t, \qquad (11.4)$$

откуда, переходя к пределу $\Delta t \to 0$, получаем

$$I = \frac{\mathcal{E} + |d\Phi/dt|}{R}, \qquad (11.5)$$

В этом случае устройство работает как генератор. Даже если мы уберем источник тока ($\mathcal{E} = 0$), в цепи все равно будет идти ток и выделяться тепло. Из уравнения (11.4) видно, что это будет происходить за счет работы внешней силы.

Обратим внимание на то, что на самом деле нет необходимости рассматривать эти случаи — электромотор и генератор — по отдельности. Легко видеть, что уравнение баланса энергии (11.2) и уравнение закона Ома (11.3) охватывают оба случая, если только учитывать знак изменения магнитного потока $\Delta\Phi$. Таким образом, одно и то же устройство может служить моделью и электромотора, и генератора. Его работа в любом режиме описывается одними и теми же уравнениями. В электрической машине направление преобразования энергии может быть изменено на обратное. Свойство обратимости присуще всем электрическим машинам постоянного тока и широко используется в технике.

Рассмотрим подробнее электромотор постоянного тока с независимым возбуждением и сопротивлением обмотки якоря R, который включен в сеть с напряжением U. Пренебрегая трением в подшипниках, исследуем условия работы электромотора: найдем зависимость тока в цепи, полной потребляемой мотором мощности и механической мощности,

развиваемой мотором, угловой скорости вращения якоря и коэффициента полезного действия от механической нагрузки, т. е. от момента внешних сил, действующих на якорь.

Предположим, что к якорю электромотора со стороны механической нагрузки приложен постоянный вращающий момент M. Будем рассматривать работу электромотора в стационарном режиме, когда угловая скорость вращения якоря ω не меняется со временем. Пусть в цепи якоря идет ток I. Воспользуемся законом сохранения энергии: потребляемую мотором от сети мощность $P=IU$ приравняем сумме механической мощности $P_м$, развиваемой мотором, и джоулева тепла, выделяющегося в якоре в единицу времени:

$$IU = P_м + I^2R. \tag{11.6}$$

Поскольку при вращении якоря в постоянном магнитном поле в его обмотке возникает э. д. с. индукции, ток I определяется законом Ома для неоднородного участка цепи:

$$I = \frac{U-\mathscr{E}}{R}, \tag{11.7}$$

где $\mathscr{E}$ — абсолютная величина э. д. с. индукции.

Сравнивая эти формулы, легко увидеть, что

$$P_м = I\mathscr{E}. \tag{11.8}$$

Это соотношение, разумеется, почти очевидно и из интуитивных соображений.

Поскольку при независимом возбуждении магнитное поле статора постоянно, то скорость изменения магнитного потока и, следовательно, э. д. с. индукции пропорциональна угловой скорости вращения якоря ω. Поэтому равенство (11.8) можно переписать в виде

$$P_м = Ik\omega, \tag{11.9}$$

где k — коэффициент пропорциональности, зависящий от индукции магнитного поля статора и от конструкции электромотора, т. е. от размеров и формы обмотки якоря. С другой стороны, развиваемая мотором механическая мощность при равномерном вращении якоря равна произведению механического момента M на угловую скорость:

$$P_м = M\omega. \tag{11.10}$$

Из сравнения выражений (11.9) и (11.10) видим, что при установившемся вращении якоря ток в его обмотке пропорционален приложенному механическому моменту:

$$I = M/k. \qquad (11.11)$$

Однако из соотношения (11.7) следует, что ток не может превышать значения

$$I_{max} = \frac{U}{R}, \qquad (11.12)$$

которое достигается при $\mathscr{E}=0$, т. е. при неподвижном якоре. Это означает, что максимальный механический момент, который может развить мотор, равен

$$M_{max} = \frac{kU}{R}. \qquad (11.13)$$

Если момент внешних сил, действующих на якорь, будет превышать эту величину, то якорь будет вращаться в противоположную сторону и мотор будет работать как генератор.

Итак, из равенства (11.11) следует, что зависимость тока в якоре от приложенного механического момента линейная (рис. 11.2, *а*), причем максимальные значения I_{max} и M_{max} даются формулами (11.12) и (11.13).

Полная мощность, потребляемая электромотором от сети, $P=IU=UM/k$, а ее максимальное значение

$$P_{max}=I_{max}U=U^2/R.$$

Зависимость потребляемой мощности от M представлена на рис. 11.2, *б* прямой *1*.

С помощью формул (11.6) и (11.11) выражение для механической мощности, развиваемой мотором, можно представить в виде

$$P_м = \frac{U}{k} M - \frac{R}{k^2} M^2. \qquad (11.14)$$

Это уравнение параболы, ветви которой направлены вниз (кривая *2* на рис. 11.2, *б*). Механическая мощность мотора обращается в нуль при $M=0$, т. е. при работе мотора на холостом ходу, и при $M=M_{max}$, когда якорь не вращается. Совсем нетрудно убедиться, что максимальное значение механической мощности мотора достигается при $M=M_{max}/2$

и равно

$$P_{\text{м max}} = \frac{1}{4} P_{\max} = \frac{1}{4} \frac{U^2}{R}.$$

Коэффициент полезного действия мотора, определяемый как отношение механической мощности к полной мощности, потребляемой от сети, равен

$$\eta = \frac{P_{\text{м}}}{P} = 1 - \frac{R}{kU} M.$$

Зависимость к. п. д. от M представлена прямой на рис. 11.2, *в*. Для определения зависимости угловой скорости якоря от механической нагрузки вернемся к формуле (11.10). Подставляя в нее $P_{\text{м}}$ из (11.14), получаем $\omega = U/k - (R/k^2) M$. График зависимости ω от M — также прямая (рис. 11.2,*г*). При отсутствии внешней нагрузки, т. е. на холостом ходу, угловая скорость максимальна:

$$\omega_0 = \frac{U}{k}. \quad (11.15)$$

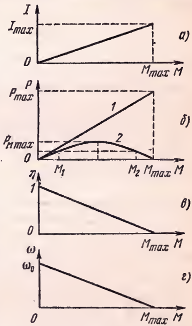


Рис. 11.2. Зависимость тока в обмотке якоря (*а*), полной и полезной мощности (*б*), к. п. д. (*в*) и угловой скорости якоря (*г*) от момента внешних сил.

С помощью формулы (11.15) коэффициент пропорциональности k между э. д. с. индукции и угловой скоростью ω, входящий в выражение для искомых величин, можно выразить через угловую скорость холостого хода ω_0, которую легко измерить на опыте.

Приведенные на рисунке кривые наглядно иллюстрируют условия работы электромотора. Стремление получить от данного электромотора возможно больший вращающий момент неизбежно приведет, как это видно из рис. 11.2, *б*, *в*,

к тому, что механическая мощность мотора и его к. п. д. будут крайне низкими, а угловая скорость якоря близка к нулю. Почти вся потребляемая при этом от сети энергия пойдет на нагревание обмотки якоря. Для получения максимальной полезной мощности от данного мотора механическую нагрузку необходимо согласовать с ним. Механический момент внешних сил должен быть равен половине максимального момента, который может развить данный электромотор. Этого можно добиться использованием редуктора. Коэффициент полезного действия мотора, работающего в таком режиме, равен 0,5. Любую механическую мощность $P_{м1}$, меньшую максимальной, можно получить при двух значениях момента внешних сил M_1 и M_2 (рис. 11.2, *б*).

Для получения высокого к. п. д. при заданной полезной мощности следует, как видно из рис. 11.2, *б*, выбирать меньшее значение момента внешних сил M_1: при этом бесполезный расход энергии на нагревание обмотки якоря будет меньше. Мотору, предназначенному для работы на разных режимах, нужен редуктор!

§ 12. Движение заряженных частиц в электрическом и магнитном полях

Как известно, сила, действующая на заряженную частицу в электромагнитном поле, имеет вид

$$\boldsymbol{F}=q\,(\boldsymbol{E}+\boldsymbol{v}\times\boldsymbol{B}). \qquad (12.1)$$

При заданных полях $\boldsymbol{E}$ и $\boldsymbol{B}$ задача о движении заряда в поле — это обычная задача классической механики о движении частицы под действием известных сил. Строго говоря, движущаяся с ускорением заряженная частица излучает электромагнитные волны и испытывает с их стороны ответное воздействие. Но этот эффект, вообще говоря, мал, и во многих случаях им можно полностью пренебречь. Но даже и тогда задача остается очень сложной, если заданные внешние поля неоднородны. В однородных электрическом и магнитном полях движение заряженной частицы происходит достаточно просто и может быть изучено элементарными методами.

Движение заряженной частицы в однородном электрическом поле совершенно аналогично движению материальной точки в однородном поле тяжести. Оно происходит с

постоянным по величине и направлению ускорением, равным произведению удельного заряда частицы q/m на напряженность поля $\boldsymbol{E}$. Траектория такого движения в общем случае представляет собой параболу. Именно так движутся

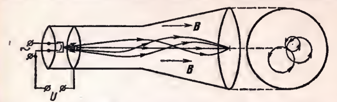


Рис. 12.1. Фокусировка пучка электронов продольным магнитным полем.

электроны в пространстве между отклоняющими пластинами в электронно-лучевой трубке осциллографа с электростатическим управлением.

Движение заряженной частицы в однородном магнитном поле под действием силы Лоренца $q\boldsymbol{v}\times\boldsymbol{B}$ происходит следующим образом. В плоскости, перпендикулярной индукции магнитного поля, частица равномерно обращается по окружности. Радиус этой окружности пропорционален величине перпендикулярной магнитному полю составляющей скорости частицы, а частота обращения от скорости не зависит и равна произведению удельного заряда частицы на индукцию магнитного поля. Если при этом частица имеет еще и составляющую скорости вдоль магнитного поля, то на такое вращение накладывается равномерное движение вдоль поля, так что траектория результирующего движения представляет собой винтовую линию. Сила Лоренца, действующая перпендикулярно скорости частицы, не меняет величины скорости и, следовательно, кинетической энергии частицы.

Интересно отметить, что при небольшом разбросе значений продольной составляющей скорости частиц движение в однородном магнитном поле обладает замечательным свойством фокусировки: выходящий из одной точки и направленный вдоль поля слегка расходящийся пучок заряженных частиц на некотором расстоянии вновь собирается в одну точку. Это свойство продольной фокусировки было использовано в 1922 году Бушем для точного измерения удельного заряда электрона.

Разберем опыт Буша подробно. Рассмотрим устройство, изображенное на рис. 12.1: электронно-лучевая трубка без управляющих пластин помещена внутрь соленоида, создающего однородное магнитное поле, направленное вдоль оси трубки. В отсутствие магнитного поля электроны летят прямолинейно и образуют на флуоресцирующем экране широкое светящееся пятно. Регулируя величину тока в соленоиде и тем самым изменяя индукцию магнитного поля, можно добиться того, что электроны соберутся на экране в яркую светящуюся точку. Выясним причину фокусировки электронов. Из электронной пушки электроны вылетают с приблизительно одинаковыми по величине скоростями, но с некоторым разбросом по направлению. Величину скорости электрона v можно определить с помощью закона сохранения энергии:

$$\frac{mv^2}{2}=eU, \qquad (12.2)$$

где e — абсолютная величина заряда электрона, а U — ускоряющее напряжение между катодом и ускоряющим анодом электронной пушки. На электрон, летящий вдоль магнитного поля, сила Лоренца не действует. Поэтому электрон, вылетевший из пушки вдоль оси трубки, движется прямолинейно и попадает в центр экрана. Если же электрон вылетел под некоторым углом α к оси трубки и, следовательно, у него есть составляющая начальной скорости, перпендикулярная магнитному полю, то, как мы видели, траектория электрона представляет собой винтовую линию: его движение есть результат сложения равномерного движения вдоль оси трубки со скоростью $v_\parallel=v\cos\alpha$ и равномерного обращения по окружности в плоскости, перпендикулярной оси трубки, со скоростью $v_\perp=v\sin\alpha$.

Угловая скорость вращения электрона по окружности определяется с помощью второго закона Ньютона:

$$\frac{mv_\perp^2}{R}=eBv_\perp, \qquad (12.3)$$

где R — радиус окружности. Учитывая связь между линейной и угловой скоростями $v_\perp=\omega_c R$, с помощью (12.3) найдем

$$\omega_c=\frac{eB}{m}, \qquad (12.4)$$

Замечательно, что угловая скорость и, следовательно, период обращения не зависят от величины скорости. Поэтому электроны, вылетевшие из пушки под разными углами, совершают полный оборот за одно и то же время. Поскольку электроны вылетают из пушки под малыми углами к оси трубки ($\cos\alpha\approx 1$), то все они движутся вдоль оси трубки практически с одной и той же скоростью $v_1\approx v$ и за время одного оборота $T=2\pi/\omega_c$ проходят вдоль оси трубки одно и то же расстояние L:

$$L=\frac{2\pi v}{\omega_c}. \qquad (12.5)$$

Это означает, что все винтовые линии, по которым движутся электроны, пересекают ось трубки практически в одной и той же точке, отстоящей на расстояние L от пушки. Такая же фокусировка происходит и после совершения электронами двух, трех и т. д. оборотов, т. е. на расстояниях $2L$, $3L$ и т. д. от пушки. Если положение одной из этих точек совпадет с плоскостью экрана, то пятно на экране сожмется в яркую точку. Разумеется, расстояние от электронной пушки до экрана определяется конструкцией трубки и не изменяется во время опыта, но мы можем изменять шаг винтовой линии L, регулируя величину индукции магнитного поля B или ускоряющее напряжение U.

Подставляя скорость электронов v из (12.2) и угловую скорость вращения ω_c из (12.4) в формулу (12.5), получаем соотношение

$$\frac{e}{m}=\frac{8\pi^2 U}{L^2B^2}. \qquad (12.6)$$

Если при неизменном ускоряющем напряжении U мы добьемся фокусировки пучка электронов, постепенно увеличивая индукцию магнитного поля B от нуля, то формула (12.6) может быть использована для вычисления отношения e/m. Для этого в правую часть нужно подставить значения U и B, при которых произошла фокусировка, а в качестве L взять расстояние от электронной пушки до экрана трубки.

Если теперь продолжать увеличивать индукцию магнитного поля, то пятно на экране будет сначала расплываться, а затем снова сожмется в яркую точку. Ясно, что теперь электроны успевают совершить два полных оборота по винтовой линии до того, как попадают на экран. Для

нахождеиия e/m в формулу (12.6) в качестве L в этом случае следует подставлять половиису расстояния от пушки до экраиа.

Отметим, что достигиутая этим методом точиость измереиия удельиого заряда электроиа составляет величину порядка десятой доли процеита.

В иастоящее время явление фокусировки пучка электроиов продольиым магиитиым полем используется во миогих электроиио-оптических приборах.

Перейдем теперь к рассмотреиию движения заряжеииой частицы в постояииых одиородиых взаимио перпеидикулярииых (так изываемых скрещеииых) электрическом и магиитном полях. Будем считать, что в иачальиый момеит частица покоится.

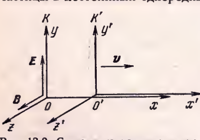


Рис. 12.2. Системы координат для изучения движеиия в скрещенных полях.

На первый взгляд кажется, что движеиие частицы будет весьма замысловатым. В самом деле, иа иеподвижиую частицу магиитиое поле ие действует, ио, как только под действием электрического поля оиа приобретает иекоторую скорость, так иемедлеино магиитиое поле будет искривлять ее траекторию. Однако, иесмотря иа кажущуюся сложность, в даииом случае удается полиостью исследовать движеиие частицы с помощью весьма простых рассуждений.

Выберем систему координат таким образом, чтобы ось z была направлеиа вдоль вектора индукции магиитиого поля $\boldsymbol{B}$, а ось y — вдоль вектора иапряжеииости электрического поля $\boldsymbol{E}$. Начало системы коордииат поместим в ту точку, где в иачальный момеит времени покоилась частица (рис. 12.2). Пусть для определеииости заряд частицы q положителеи.

Прежде всего убедимся, что траектория представляет собой плоскую кривую. Первоначальио покоившейся частице электрическое поле сообщает ускореиие и, следовательно, скорость вдоль оси y. Поскольку сила, действующая на частицу со стороны магнитного поля, перпеидикулярна как индукции поля, так и скоростн частицы, то и эта сила также действует в плоскости xy. Другими словами, ускоре-

ние частицы, а следовательно, и скорость вдоль оси z равны нулю: частица никогда не сможет покинуть плоскость xy. Но и в плоскости xy первоначально покоившаяся положительно заряженная частица может двигаться только в верхней полуплоскости ($y \geqslant 0$). В этом проще всего убедиться из энергетических соображений. В самом деле, постоянное магнитное поле, действуя перпендикулярно скорости, работы не совершает, а постоянное электрическое поле потенциально. В рассматриваемом однородном электрическом поле потенциальная энергия заряженной частицы зависит только от координаты y, и наша частица, оказавшись ниже оси x, имела бы полную энергию бо́льшую, чем в начальный момент. Самое большее — частица сможет только дойти до оси x, но при этом скорость ее должна обратиться в нуль.

Чтобы продвинуться дальше в выяснении вопроса о форме траектории, забудем на время о начальных условиях и задумаемся над таким вопросом: может ли заряженная частица в скрещенных электрическом и магнитном полях двигаться с постоянной скоростью? Очевидно, что для этого полная сила, действующая на частицу, должна быть равна нулю, т. е. магнитная и электрическая силы должны быть равны по величине и противоположны по направлению. Электрическая сила, действующая на положительно заряженную частицу, направлена вдоль оси y, следовательно, магнитная должна быть направлена в отрицательном направлении этой оси. Нетрудно убедиться, что для этого скорость частицы должна быть направлена вдоль оси x. Величина скорости определяется из соотношения

$$qE = qvB, \tag{12.7}$$

откуда

$$v = \frac{E}{B}. \tag{12.8}$$

Поскольку скорость частицы не может превышать скорости света в вакууме c, то из формулы (12.8) видно, что движение заряженной частицы в скрещенных полях с постоянной скоростью возможно только при $E < cB$. В противном случае условие (12.7) не может быть выполнено.

Рассмотрим поведение частицы, движущейся в скрещенных полях с постоянной скоростью, с точки зрения системы отсчета, в которой эта частица покоится (рис. 12.2). Поскольку такая система отсчета K' движется относительно

исходной системы K равномерно и прямолинейно, то она также является инерциальной. Полная сила (12.1), действующая на покоящуюся в этой системе отсчета частицу, должна быть равна нулю. Но магнитная сила отсутствует, так как частица покоится. Следовательно, должна отсутствовать и электрическая сила. Но это возможно, только если напряженность электрического поля E' в системе K' равна нулю.

Таким образом, мы приходим к выводу, что в системе отсчета K', движущейся относительно исходной системы K с постоянной скоростью $v=E/B$ вдоль оси x, электрическое поле отсутствует и есть только магнитное поле. Этот пример еще раз иллюстрирует относительный характер электромагнитного поля: напряженность электрического поля и индукция магнитного поля изменяются при переходе от одной инерциальной системы отсчета к другой.

Разумеется, мы могли прийти к выводу об отсутствии электрического поля в системе K' и непосредственно с помощью формул преобразования полей (10.5), если с самого начала ограничиться случаем, когда движение частицы происходит со скоростью, малой по сравнению со скоростью света. Вторая из формул (10.5) показывает, что в рассматриваемом случае при $v \ll c$ индукция магнитного поля при переходе в систему отсчета K' практически не меняется: $\mathbf{B}'=\mathbf{B}$.

Теперь возвратимся к рассмотрению первоначально покоившейся в системе отсчета K частицы и рассмотрим ее движение с точки зрения наблюдателя в системе отсчета K', где есть только магнитное поле. Очевидно, что начальная скорость этой частицы в системе отсчета K' направлена в отрицательном направлении оси x и в соответствии с формулой (12.8) равна $v=E/B$. Как мы видели в предыдущем примере, в однородном магнитном поле заряженная частица равномерно обращается по окружности, если ее скорость перпендикулярна магнитному полю. Радиус окружности R и угловая скорость ω_c даются соотношениями

$$R=\frac{v}{\omega_c}, \quad \omega_c=\frac{qB}{m}. \tag{12.9}$$

Положение окружности, по которой движется частица в системе отсчета K', показано на рис. 12.3. Частица с положительным зарядом обращается по часовой стрелке.

Найдем теперь движение частицы в системе отсчета K. Очевидио, что оио получается в результате сложения равиомерного движения точки по окружиости и поступательного движеиия этой окружиости со скоростью v по оси x. Легко

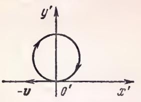


Рис. 12.3. В системе отсчета K' заряд движется по окружиости.

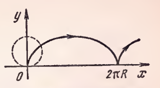


Рис. 12.4. В системе отсчета K траектория заряда — циклоида.

сообразить, что точио так же движется та точка обода колеса радиуса R, катящегося без проскальзывания по оси x, которая в начальный момент находилась в начале координат. Траектория такого движеиия носит иазвание циклоиды. Она изображеиа на рис. 12.4.

Нетрудно получить зависимость координат частицы от времеии. Для этого нужно просто «прокатить» колесо из начальиого положеиия по оси x (рис. 12.5) и выразить координаты интересующей иас точки A через угол поворота колеса $\varphi=\omega_c t$. При качении без проскальзывания длииа дуги AD равна длине отрезка OD, поэтому

$$x(t)=R(\omega_c t-\sin\omega_c t),$$
$$y(t)=R(1-\cos\omega_c t),$$

где R и ω_c определяются формулами (12.9).

Приведениое рассмотреиие справедливо и для частицы с отрицательным зарядом. Такая частица в магнитном поле обращается по окружиости против часовой стрелки, ио центр окружиости перемещается по-прежиему в направлении оси x. Траектория в этом случае является зеркальным отражением в плоскости $y=0$ кривой на рис. 12.4.

Таким образом, любая частица, иезависимо от ее массы, величины и знака заряда, в скрещенных полях совершает дрейф в иаправлеиии, перпендикулярном к векторам $\boldsymbol{E}$ и $\boldsymbol{B}$, с одной и той же скоростью $v=E/B$. Хотя этот результат получен нами для первоначально покоившейся частицы, ои

имеет совершенно общий характер. В самом деле, если частица имеет какую-то начальную скорость в плоскости xy, то это скажется только на положении центра и величине радиуса окружности, по которой частица обращается в системе отсчета K'. Центр этой окружности в системе отсчета K во всех случаях равномерно движется (дрейфует)

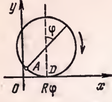


Рис. 12.5. К выводу уравнения циклоиды.

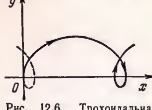


Рис. 12.6. Трохоидальная траектория.

с той же самой скоростью $v=E/B$. Результирующее движение частицы будет складываться из равномерного движения вдоль оси x и равномерного вращения по окружности в плоскости xy. Траектория этого движения носит название трохоиды. В общем случае это не соответствует качению без проскальзывания. Один из возможных типов трохоиды показан на рис. 12.6. В частном случае трохоидальная траектория может выродиться даже в прямую линию, параллельную оси x. Легко сообразить, что так будет, если начальная скорость частицы направлена вдоль оси x и равна E/B, так как именно в этом случае в системе отсчета K' частица покоится.

В заключение отметим, что если у начальной скорости частицы есть составляющая вдоль магнитного поля (т. е. вдоль оси z), то на уже рассмотренное движение наложится еще и равномерное движение в этом направлении.

ПЕРЕМЕННЫЙ ЭЛЕКТРИЧЕСКИЙ ТОК

§ 13. Цепи переменного тока. Векторные диаграммы. Резонанс

Физические процессы, происходящие в цепях синусоидального переменного тока, представляют собой вынужденные колебания, полный анализ которых сводится к решению дифференциальных уравнений с постоянными коэффициен-

тами. Одиако наиболее важный для практических примеиеиий устаиовившийся режим таких колебаний, когда собствеиные колебаиия в цепи уже затухли, может быть строго исследоваи элемеитариыми методами.

Прежде всего рассмотрим простейшие случаи, когда переменное напряжение $U(t)=U_0\cos\omega t$ подается иа нагрузку, представляющую собой либо обычное омическое сопротивлеиие R, либо емкость C, либо иидуктивиость L.

В случае активного сопротивлеиия R ток в цепи I определяется соотиошеиием

$$I(t)=\frac{U(t)}{R}=I_0\cos\omega t,\quad I_0=\frac{U_0}{R}, \qquad (13.1)$$

откуда видио, что в такой цепи ие происходит сдвига по фазе между иапряжением и током.

В цепи, содержащей только емкость C, ток проще всего найти, воспользовавшись тем, что его величииа определяется скоростью изменеиия заряда коиденсатора: $I=dq/dt$. Так как $q=CU$, а емкость конденсатора постояииа, то для тока получаем

$$I(t)=-C\omega U_0\sin\omega t=C\omega U_0\cos\left(\omega t+\frac{\pi}{2}\right). \qquad (13.2)$$

Таким образом, ток в цепи имеет сииусоидальиый характер и опережает по фазе напряжеиие иа $\pi/2$:

$$I(t)=I_0\cos\left(\omega t+\frac{\pi}{2}\right).$$

Связи между амплитудными значеииями подаваемого иапряжения U_0 и тока в цепи I_0 можио, как видио из (13.2), придать вид закоиа Ома, если ввести поиятие зависящего от частоты ω емкостиого сопротивлеиия R_C:

$$I_0=\frac{U_0}{R_C},\quad R_C=\frac{1}{\omega C}. \qquad (13.3)$$

Полученный результат можно наглядио проиллюстрировать с помощью графиков зависимости напряжения и тока от времени (рис. 13.1). В те момеиты времеии, когда подаваемое напряжение достигает экстремальных значеиий, заряд иа коидеисаторе ие меияется и, следовательно, ток в цепи обращается в нуль. В точках, где напряжение обращается в нуль, величина его меняется наиболее быстро и,

следовательно, ток достигает экстремальных значений. Итак, физическая причина сдвига по фазе очевидна, величина сдвига равна $\pi/2$, а направление сдвига (опережение или отставание по фазе) легко установить, рассматривая, например, первую четверть периода изменения напряжения: напряжение убывает, т. е. конденсатор разряжается, несмотря на то, что ток увеличивается по абсолютной величине. Это возможно, только если напряжение и ток имеют противоположные знаки, т. е. график тока действительно имеет вид, изображенный на рис. 13.1.

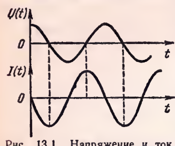


Рис. 13.1. Напряжение и ток в конденсаторе.

Случай, когда синусоидальное напряжение подается на индуктивность L, проще всего проанализировать, сравнивая выражения

$$I = C\frac{dU}{dt}, \quad U = L\frac{dI}{dt}. \tag{13.4}$$

Первая формула представляет собой выражение для тока в только что рассмотренной цепи, содержащей емкость C. Второе соотношение отражает тот факт, что поданное на индуктивность L синусоидальное напряжение U в каждый момент времени компенсирует возникающую в катушке электродвижущую силу самоиндукции $\mathcal{E} = -L\,dI/dt$. Анализ первого из соотношений (13.4) привел к формуле (13.2). Следовательно, формула такого же типа будет получена при анализе второго из соотношений (13.4). Она получается из (13.2) заменой $I \rightleftarrows U$, $C \to L$:

$$U(t) = L\omega I_0 \cos\left(\omega t + \frac{\pi}{2}\right). \tag{13.5}$$

Из первого соотношения (13.4) следует, что ток I опережает напряжение U на $\pi/2$; аналогично, из второго следует, что в такой цепи напряжение опережает ток на $\pi/2$. Задаваемой величиной является подаваемое напряжение $U(t) = U_0 \cos \omega t$, поэтому для тока I из (13.5) получаем

$$I(t) = I_0 \cos\left(\omega t - \frac{\pi}{2}\right).$$

Как и раньше, связи между амплитудными значениями тока и напряжения можно придать вид закона Ома, если, воспользовавшись (13.5), ввести индуктивное сопротивление R_L:

$$I_0=\frac{U_0}{R_L}, \quad R_L=\omega L. \tag{13.6}$$

Полученный результат также можно проиллюстрировать с помощью графиков (рис. 13.2). На верхнем графике показана зависимость тока от времени. На втором графике изображена э. д. с. самоиндукции. Положение экстремумов и сдвиг этого графика относительно графика тока легко определить с помощью закона электромагнитной индукции и правила Ленца: $\mathscr{E}=-L\,dI/dt$. Действительно, э. д. с. индукции обращается в нуль в точках экстремума тока и достигает экстремальных значений в те моменты, когда ток меняется наиболее быстро. В каждый момент полярность э. д. с. самоиндукции такова, чтобы препятствовать изменению тока,— этим сразу устанавливается направление сдвига по фазе между током и э. д. с. самоиндукции. И, наконец, приложенное напряжение изменяется в противофазе с э. д. с. самоиндукции (нижний график на рис. 13.2).

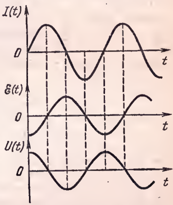


Рис. 13.2. Ток, э. д. с. самоиндукции и напряжение на индуктивности.

Рассмотрение этих простейших цепей показывает, что, за исключением случая активного сопротивления R, невозможно написать закон Ома для цепей переменного тока, определяющий мгновенное значение тока $I(t)$ в виде отношения приложенного напряжения к сопротивлению соответствующего участка, вследствие того, что между током и напряжением существует сдвиг по фазе. Как мы видели, закон Ома справедлив только для амплитудных значений тока и напряжения.

Теперь рассмотрим последовательную цепь переменного тока, содержащую активное сопротивление R, емкость C и

индуктивность L (рис. 13.3). К этой цепи приложено напряжение $U(t)=U_0 \cos \omega t$.

В последовательной цепи переменного тока сила тока I в каждый момент времени во всех участках цепи одинакова, а сумма мгновенных значений напряжений на сопротивлении R, емкости C и индуктивности L равна значению приложенного напряжения в тот же момент времени:

$$U=U_R+U_C+U_L. \qquad (13.7)$$

Воспользуемся тем обстоятельством, что мгновенное значение любой изменяющейся по гармоническому закону величины можно представить как проекцию вектора на некоторое заранее выбранное направление, причем сам вектор равномерно вращается в плоскости с частотой ω, а его длина равна амплитудному значению исследуемой величины. С помощью такого представления исследуемой схеме можно сопоставить векторную диаграмму, изображенную на рис. 13.4, *а*. Каждой величине: току I, напряжениям на сопротивлении R, емкости C и индуктивности L — сопоставляются векторы, длина которых равна амплитудному значению соответствующей величины. Вся система векторов вращается как целое с угловой скоростью ω вокруг оси, перпендикулярной плоскости рисунка. Мгновенные значения величин I, U_R, U_L и U_C получаются проектированием соответствующих векторов на заранее выбранное направление NN. Поскольку, как мы видели, ток в цепи находится в фазе с напряжением U_R, отстает на $\pi/2$ от напряжения на индуктивности U_L и опережает на $\pi/2$ напряжение на емкости U_C, то при выбранном направлении вращения вектор U_{0L} опережает векторы I_0 и U_{0R} на $\pi/2$, которые в свою очередь опережают на $\pi/2$ вектор U_{0C}.

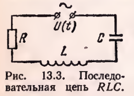


Рис. 13.3. Последовательная цепь RLC.

Как теперь найти вектор U_0, изображающий приложенное напряжение U? Легко видеть, что для этого нужно просто найти сумму векторов U_{0R}, U_{0L} и U_{0C}, так как проекция результирующего вектора, которая и определяет мгновенное значение приложенного напряжения U, равна сумме проекций составляющих векторов, представляющих собой мгновенные значения напряжений U_R, U_L и U_C, в полном соответствии с равенством (13.7) (рис. 13.4, *б*). Из этого

рисунка легко видеть, что

$$U_0^2 = U_{0R}^2 + (U_{0L} - U_{0C})^2, \quad \operatorname{tg}\varphi = \frac{U_{0L} - U_{0C}}{U_{0R}}. \quad (13.8)$$

Используя связь между амплитудным значением тока I_0 и

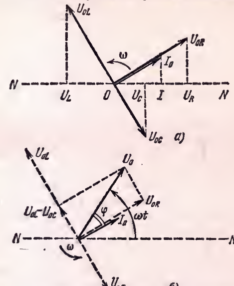


Рис. 13.4. *а*) Векторная диаграмма для последовательной цепи *RLC* (рис. 13.3). *б*) К определению связи между приложенным напряжением и током в цепи.

амплитудными значениями напряжений на отдельных элементах цепи:

$$U_{0R} = I_0 R, \quad U_{0C} = \frac{I_0}{\omega C}, \quad U_{0L} = I_0 \omega L,$$

с помощью (13.8) получаем

$$I_0 = \frac{U_0}{\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}}, \quad (13.9)$$

$$\operatorname{tg}\varphi = \frac{\omega L - \frac{1}{\omega C}}{R}. \quad (13.10)$$

Итак, если приложенное напряжение $U(t) = U_0 \cos \omega t$, то ток в цепи $I(t) = I_0 \cos(\omega t - \varphi)$, где I_0 и φ определяются

формулами (13.9) и (13.10). Ток в цепи, как и напряжение, меняется по синусоидальному закону, но между током и напряжением существует сдвиг по фазе, равный φ.

С помощью векторной диаграммы на рис. 13.4, *б* теперь легко написать выражения для мгновенных значений напряжений на отдельных элементах схемы:

$$U_R = I_0 R \cos(\omega t - \varphi),$$
$$U_L = I_0 \omega L \cos\left(\omega t - \varphi + \frac{\pi}{2}\right),$$
$$U_C = \frac{I_0}{\omega C} \cos\left(\omega t - \varphi - \frac{\pi}{2}\right).$$

Выясним, что покажет вольтметр, если его подключить к какому-либо из элементов схемы. Если вольтметр проградуирован так, что он показывает действующее значение напряжения, то его показания будут в $\sqrt{2}$ раз меньше амплитудного значения напряжения на том элементе схемы, к которому он подключен. Произведя измерения напряжений на всех элементах схемы по отдельности, можно убедиться, что сумма этих напряжений всегда больше действующего значения подаваемого на схему напряжения. Более того, напряжение на любом из реактивных сопротивлений может быть гораздо больше подаваемого напряжения. Напряжение же на активном сопротивлении никогда не бывает больше подаваемого напряжения.

Если при измерении напряжений на реактивных элементах напряжения окажутся равными друг другу, то это значит, что равны реактивные сопротивления: $\omega L = 1/\omega C$. Такую ситуацию называют резонансом напряжений в цепи переменного тока. При этом напряжение на активном сопротивлении равно приложенному внешнему напряжению. Сопротивление всей последовательной цепи при резонансе напряжений становится чисто активным и равным R. Сдвиг фаз между приложенным напряжением и током в этом случае отсутствует.

Перейдем теперь к рассмотрению цепи переменного тока, содержащей параллельно соединенные активное сопротивление R, индуктивность L и емкость C (рис. 13.5), на которую подается переменное синусоидальное напряжение $U(t) = U_0 \cos \omega t$. Как и в случае последовательного соединения элементов, эту цепь удобно исследовать с помощью вектор-

ных диаграмм. Напряжение на всех параллельно соединенных элементах одинаково и равно приложенному напряжению $U(t)$. Мгновенное значение тока в неразветвленной части цепи $I(t)$ равно алгебраической сумме токов в параллельных участках:

$$I = I_R + I_C + I_L. \qquad (13.11)$$

Поскольку ток через сопротивление R находится в фазе с приложенным напряжением, ток в ветви, содержащей емкость, опережает напряжение на $\pi/2$, а ток через индуктивность отстает от напряжения на $\pi/2$, то векторная диаграмма, соответствующая этой цепи, имеет вид, изображенный на рис. 13.6. Учитывая связь между амплитудными значениями токов в различных элементах и амплитудным значением приложенного напряжения:

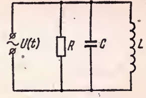


Рис. 13.5. Параллельное соединение R, L и C.

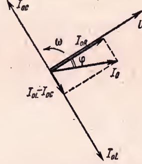


Рис. 13.6. Векторная диаграмма для параллельной цепи RLC (рис. 13.5).

$$U_0 = I_{0R}R = I_{0C}\frac{1}{\omega C} = I_{0L}\omega L,$$

с помощью векторной диаграммы на рис. 13.6 нетрудно получить следующие выражения для амплитуды тока в неразветвленной части цепи и для сдвига по фазе между приложенным напряжением и этим током:

$$I_0 = U_0\sqrt{\frac{1}{R^2} + \left(\frac{1}{\omega L} - \omega C\right)^2}, \qquad (13.12)$$

$$\operatorname{tg}\varphi = R\left(\frac{1}{\omega L} - \omega C\right). \qquad (13.13)$$

Таким образом, ток в неразветвленной части цепи равен $I(t) = I_0\cos(\omega t - \varphi)$, где I_0 и φ определяются формулами

(13.12), (13.13). Векторная диаграмма дает также возможность написать выражения для мгновенных значений тока в отдельных ветвях цепи:

$$I_R = \frac{U_0}{R}\cos\omega t,$$

$$I_L = \frac{U_0}{\omega L}\cos\left(\omega t - \frac{\pi}{2}\right),$$

$$I_C = U_0\,\omega C\cos\left(\omega t + \frac{\pi}{2}\right).$$

При равенстве емкостного и индуктивного сопротивлений, т. е. при $\omega L = 1/\omega C$, сдвиг фаз между током в неразветвленной части цепи и напряжением обращается в нуль. Токи I_C и I_L при этом равны по величине, и так как они находятся в противофазе, то ток в неразветвленной части становится равным току I_R через активное сопротивление. Заметим, что токи I_L и I_C в отдельных ветвях цепи могут значительно превосходить ток в подводящих проводах. Такая ситуация носит название резонанса токов. При этом происходит обмен энергией между электрическим и магнитным полями, сосредоточенными в емкости и индуктивности, а источник питания только компенсирует потери энергии за счет выделения джоулева тепла на сопротивлении R. Если сопротивление R вообще убрать из цепи ($R \to \infty$), то энергетические потери в такой идеализированной схеме вообще отсутствуют и ток в подводящих проводах равен нулю, хотя в контуре, состоящем из L и C, ток может быть очень большим. В этом случае на резонансной частоте $\omega = 1/\sqrt{LC}$ полное сопротивление контура неограниченно возрастает.

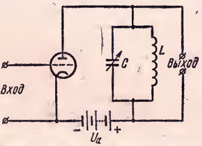


Рис. 13.7. Принципиальная схема резонансного усилителя.

Резонанс токов, наряду с резонансом напряжений, широко используется в технике. В качестве примера рассмотрим работу простейшего резонансного усилителя, в анодной цепи которого имеется колебательный контур, настраи-

ваемый на частоту сигнала, который требуется усилить (рис. 13.7). Для резонансной частоты (точнее, для узкой полосы частот) контур представляет собой большое сопротивление, и резонансный усилитель действует таким же образом, как и обычный усилитель на сопротивлениях, причем роль анодного сопротивления играет колебательный контур. Для частот, заметно отличающихся от резонансной, контур представляет собой практически короткое замыкание анодной цепи, и поэтому усиления сигнала не происходит.

Другим важным примером использования резонанса токов является индукционная печь, в которой нагревание металлов производится вихревыми токами. Параллельно нагревающей катушке присоединяют конденсатор и подбирают его емкость так, чтобы получить на частоте питающего генератора резонанс токов. Тогда через подводящие провода и генератор пойдет сравнительно небольшой ток, который может быть во много раз меньше тока в колебательном контуре, образованном конденсатором и нагревающей катушкой.

§ 14. Мощность переменного тока. Преобразование и передача электроэнергии. Трансформатор

Широкое использование переменного тока в народном хозяйстве связано с удобством его преобразования с помощью трансформаторов и исключительной простотой повсеместно применяемых асинхронных двигателей. Но почему из всех возможных форм периодических переменных токов наибольшее распространение получили переменные токи синусоидальной формы? Дело в том, что синусоидальные токи по сравнению со всеми другими токами позволяют наиболее просто и экономично осуществлять передачу, распределение, преобразование и использование электрической энергии. Только при помощи синусоидальных токов удается сохранить неизменными формы кривых напряжений и токов на всех участках линейной электрической цепи, т. е. цепи, содержащей резисторы, конденсаторы, катушки индуктивности, но не содержащей нелинейных элементов — диодов, транзисторов, электронных ламп и т. п. Чтобы убедиться в этом, рассмотрим следующую простую линейную цепь, состоящую из конденсатора C и сопротивления R

(рис. 14.1). Будем подавать на вход этой цепи переменное напряжение различной формы и смотреть, какую форму будет иметь выходное напряжение, равное напряжению на сопротивлении R. Если подать на вход синусоидальное напряжение $U_{вх}(t) = U_0 \cos \omega t$, то через последовательно соединенные конденсатор C и сопротивление R пойдет синусоидальный переменный ток, опережающий по фазе приложенное напряжение $U_{вх}$ на угол φ, как это видно из векторной диаграммы на рис. 14.2,*а*. Напряжение на сопротивлении R будет также синусоидальным и в фазе с током. Таким образом, выходное напряжение $U_{вых}(t)$, как и входное, будет синусоидальным, но сдвинутым относительно него по фазе на угол φ (рис. 14.2, *б*).

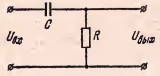


Рис. 14.1. RC-цепочка.

Посмотрим теперь, что будет на выходе этой цепи, если на ее вход подавать напряжение в виде прямоугольных

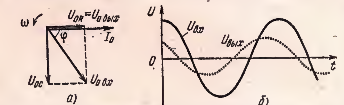


Рис. 14.2. Векторная диаграмма для RC-цепочки (*а*) и графики входного и выходного напряжений (*б*).

импульсов (рис. 14.3). Начало каждого прямоугольного импульса соответствует подключению к цепи источника постоянного напряжения на время, равное длительности импульса. При этом в цепи скачком возникает ток, который постепенно уменьшается по мере того, как конденсатор заряжается. Время, в течение которого продолжается процесс заряда конденсатора, зависит от величины RC. Если это время меньше длительности подаваемого на вход прямоугольного импульса, то ток заряда прекратится раньше, чем закончится прямоугольный импульс. Именно этот случай изображен на рис. 14.3. В момент прихода заднего фронта прямоугольного импульса подаваемое напряжение

скачком обращается в нуль. Но этого можно добиться только путем короткого замыкания входных клемм схемы. Цепь, содержащая R и C, становится короткозамкнутой, и конденсатор C разряжается через сопротивление R. Направление тока разряда противоположно зарядному току, поэтому выходное напряжение на сопротивлении R имеет противоположную полярность (рис. 14.3). Таким образом, форма выходного напряжения оказывается совершенно иной, чем форма входного напряжения.

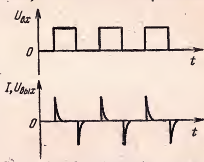


Рис. 14.3. Преобразование прямоугольных импульсов напряжения RC-цепочкой.

Итак, для сохранения формы передаваемого напряжения необходимо использовать именно синусоидальный переменный ток. Но для такого тока при наличии в цепи реактивных элементов между напряжением и током возникает сдвиг по фазе на некоторый угол φ. Это, как известно, приводит к появлению множителя $\cos\varphi$ в выражении для мощности P, развиваемой переменным током на нагрузке:

$$P = \frac{1}{2} U_0 I_0 \cos\varphi = UI \cos\varphi, \qquad (14.1)$$

где U и I — действующие значения напряжения и тока, в $\sqrt{2}$ раз меньшие амплитудных значений. Потребителю обычно подается напряжение определенной величины U, поэтому одна и та же мощность P будет потребляться при разных значениях тока в цепи I в зависимости от величины сдвига фазы между током и напряжением. При малых значениях $\cos\varphi$ ток должен быть большим, что приводит к большим тепловым потерям в подводящих проводах линии передачи. Если r — сопротивление линии передачи, то рассеиваемая мощность тепловых потерь в линии P_1 равна $I^2 r$. Выражая ток в цепи с помощью (14.1), для P_1 получим

$$P_1 = \frac{P^2}{U^2 \cos^2\varphi} r. \qquad (14.2)$$

Для уменьшения потерь следует добиваться как можно меньшего сдвига фазы между током и напряжением в нагрузке.

Большинство современных потребителей электрической энергии синусоидального тока представляют собой нагрузки индуктивного характера, токи в которых отстают по фазе

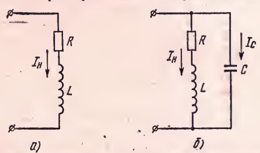


Рис. 14.4. Эквивалентная схема потребителя с индуктивной нагрузкой (*а*) и включение вспомогательного конденсатора для увеличения $\cos\varphi$ (*б*).

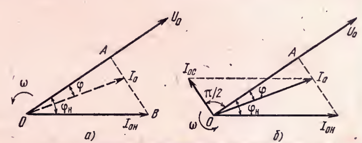


Рис. 14.5. Векторные диаграммы для цепей на рис. 14.4.

от напряжения источника питания. Эквивалентную схему такого потребителя можно изобразить в виде последовательно соединенных активного сопротивления R и индуктивности L (рис. 14.4, *а*). Соответствующая векторная диаграмма показана на рис. 14.5, *а*. Ток $I_н$ через нагрузку отстает от приложенного напряжения на определенный угол $\varphi_н$. Потребляемая нагрузкой мощность согласно (14.1) равна

$$P = \frac{1}{2} U_0 I_{0н} \cos\varphi_н. \qquad (14.3)$$

Из этой формулы видно, что при напряжении U_0 такую же мощность можно было бы получить и при любом другом токе I_0 таком, что изображающий его вектор (показанный пунктиром на рис. 14.5, *а*) оканчивается на перпендикуляре AB, опущенном из конца $I_{0н}$ на направление U_0, так как при этом $I_0 \cos\varphi = I_{0н} \cos\varphi_н$. Но если $\varphi < \varphi_н$, то $I_0 < I_{0н}$ и при той же мощности тепловые потери в подводящих проводах будут меньше. Как же добиться того, чтобы сдвиг фаз между напряжением и током в цепи уменьшился? Легко сообразить, что для этого можно подсоединить параллельно нагрузке вспомогательный конденсатор (рис. 14.4, *б*). Векторная диаграмма в этом случае будет иметь вид, изображенный на рис. 14.5, *б*. Векторы, изображающие приложенное напряжение U и ток через нагрузку $I_н$, останутся неизменными, а полный ток в неразветвленной цепи, равный сумме токов через нагрузку и вспомогательный конденсатор, будет изображаться вектором I_0. Подбирая величину емкости конденсатора, можно добиться любого, в том числе и нулевого, сдвига по фазе между приложенным напряжением и током в неразветвленной части цепи. Вычислим, какая необходима емкость для того, чтобы сдвиг по фазе принял заданное значение φ. Из рис. 14.5, *б* видно, что длина вектора I_{0C} равна

$$I_{0C} = OA\,(\operatorname{tg}\varphi_н - \operatorname{tg}\varphi). \qquad (14.4)$$

Но $OA = I_{0н}\cos\varphi_н$, и с помощью (14.3) находим $OA = 2P/U_0$. Амплитудное значение тока в конденсаторе I_{0C} связано с амплитудным значением подаваемого напряжения формулой $I_{0C} = U_0\omega C$. Подставляя OA и I_{0C} в (14.4), находим

$$C = \frac{2P}{U_0^2\,\omega}\,(\operatorname{tg}\varphi_н - \operatorname{tg}\varphi). \qquad (14.5)$$

Таким образом, существует достаточно простой и эффективный способ снижения потерь в линиях передачи энергии переменного тока, связанных с реактивным характером сопротивления нагрузки. Но даже в том случае, когда сопротивление нагрузки является чисто активным и сдвиг фаз между напряжением и током отсутствует, т. е. $\cos\varphi = 1$, тепловые потери в линии передачи все равно неизбежны. Можно ли их каким-либо способом уменьшить? Ответ на этот вопрос дает формула (14.2). Из нее видно, что при заданном значении передаваемой потребителю мощности P

уменьшить тепловые потери в линии можно, либо уменьшая сопротивление r проводов линии передачи, либо повышая напряжение U переменного тока, подаваемого потребителю. Уменьшение сопротивления линии в настоящее время возможно лишь до известных пределов, поэтому до создания эффективных сверхпроводящих линий электропередачи с потерями приходится бороться повышением напряжения.

Для преобразования напряжения на электростанциях и у потребителей используются трансформаторы. Рассмотрим принцип действия трансформатора. Пусть сначала вторичная обмотка трансформатора разомкнута, а на первичную подается переменное синусоидальное напряжение. Это режим холостого хода. Как и всякую катушку индуктивности, первичную обмотку трансформатора можно рассматривать как последовательно соединенные индуктивность L и активное сопротивление R. Напряжение U_L на индуктивном сопротивлении $R_L = \omega L$ первичной обмотки опережает по фазе ток и, следовательно, напряжение U_R на ее активном сопротивлении на угол, равный $\pi/2$. Поэтому амплитудные значения поданного на первичную обмотку напряжения U_{01} и напряжений на R и L связаны соотношением

$$U_{01} = \sqrt{U_{0R}^2 + U_{0L}^2}. \tag{14.6}$$

Разумеется, непосредственно измерить U_L и U_R по отдельности невозможно, так как первичная обмотка, строго говоря, не есть последовательно соединенные индуктивность L и активное сопротивление R: каждый элемент обмотки обладает одновременно индуктивностью и сопротивлением. Это так называемая цепь с распределенными параметрами. Но при расчете можно заменить реальную обмотку на цепь с сосредоточенными параметрами — катушку индуктивности и сопротивление, соединенные последовательно, поскольку через каждый элемент исходной цепи идет один и тот же ток.

Напряжение на индуктивности U_L в каждый момент времени компенсирует возникающую в первичной обмотке э. д. с. самоиндукции $\mathcal{E}_1$, поэтому

$$U_L = -\mathcal{E}_1. \tag{14.7}$$

Если весь магнитный поток, создаваемый током первичной обмотки, целиком, т. е. без рассеяния, пронизывает вторич-

ную обмотку, то индуцируемая в каждом витке вторичной обмотки э. д. с. будет такой же, как и в каждом витке первичной обмотки. Поэтому отношение электродвижущих сил в первичной и вторичной обмотках равно отношению чисел витков:

$$\frac{\mathscr{E}_1}{\mathscr{E}_2}=\frac{n_1}{n_2}. \qquad (14.8)$$

На выходе разомкнутой вторичной обмотки существует напряжение, равное индуцируемой в ней э. д. с.:

$$U_2=\mathscr{E}_2. \qquad (14.9)$$

Подставляя сюда $\mathscr{E}_2$ из (14.8) и учитывая (14.7), получим

$$U_2=\frac{n_2}{n_1}\mathscr{E}_1=-\frac{n_2}{n_1}U_L. \qquad (14.10)$$

Таким образом, величина напряжения на разомкнутой вторичной обмотке трансформатора пропорциональна не подаваемому на первичную обмотку напряжению U_1, а лишь напряжению на индуктивном сопротивлении первичной обмотки U_L. Отсюда сразу становится ясна роль сердечника трансформатора. В самом деле, из формулы (14.6) следует, что напряжение на индуктивности U_L будет тем ближе к подаваемому на вход трансформатора напряжению U_1, чем больше будет индуктивное сопротивление первичной обмотки ωL по сравнению с ее активным сопротивлением R. Наличие сердечника из материала с высокой магнитной проницаемостью приводит к многократному увеличению индуктивности L. У такого трансформатора на холостом ходу $U_2\approx-\frac{n_2}{n_1}U_1$. Знак минус означает, что эти напряжения находятся в противофазе. Благодаря большому индуктивному сопротивлению первичной обмотки ток в ней при разомкнутой вторичной цепи мал.

При замыкании вторичной цепи трансформатора на некоторую нагрузку во вторичной обмотке появляется ток. Создаваемый этим током магнитный поток направлен так, что, согласно правилу Ленца, препятствует изменению магнитного потока, создаваемого током в первичной обмотке. Если бы при этом ток в первичной обмотке остался неизменным, то это привело бы к уменьшению магнитного потока. Значит, включение нагрузки во вторичную цепь эквивалент-

но уменьшению индуктивности первичной цепи. Но уменьшение индуктивного сопротивления немедленно приводит к увеличению тока в первичной обмотке, к уменьшению сдвига по фазе между напряжением и током и, следовательно, к увеличению потребляемой от внешней цепи мощности. Таким образом, если на холостом ходу трансформатор представляет собой почти чисто индуктивное сопротивление, то по мере увеличения нагрузки трансформатора, т. е. тока во вторичной цепи, характер сопротивления трансформатора становится все ближе к активному.

Если потери энергии в самом трансформаторе малы, то на основании закона сохранения энергии потребляемая трансформатором мощность целиком передается нагрузке. Тогда с помощью (14.1) можно написать

$$\frac{1}{2} U_{01} I_{01} \cos \varphi_1 = \frac{1}{2} U_{02} I_{02} \cos \varphi_2, \qquad (14.11)$$

где φ_1 и φ_2 — сдвиги фаз между током и напряжением в первичной и вторичной цепях.

Приведенное выше рассмотрение работы трансформатора относится к идеализированному случаю трансформатора без потерь. В реальном трансформаторе всегда имеются потери, связанные с выделением джоулева тепла в обмотках, с токами Фуко, с работой перемагничивания, обусловленной гистерезисом в сердечнике, и с рассеянием магнитного потока. Но в современных трансформаторах суммарные потери не превышают нескольких процентов от передаваемой мощности. Коэффициент полезного действия трансформаторов очень высок и лежит в пределах 95—99,5%.

§ 15. Трехфазный ток. Электрические машины переменного тока

Наряду с простым синусоидальным переменным током в технике широко используется так называемый трехфазный ток.

Представим себе прямоугольную проволочную рамку с несколькими витками, равномерно вращающуюся в однородном магнитном поле. Возникающая в этой рамке э. д. с. индукции меняется по синусоидальному закону. Если же вокруг общей оси вращается не одна, а три одинаковые рамки, плоскости которых повернуты друг относительно друга

на 120°, то возникающие в них синусоидальные э. д. с. будут сдвинуты по фазе на 120° (рис. 15.1):

$$\mathscr{E}_1 = \mathscr{E}_0 \sin \omega t,$$
$$\mathscr{E}_2 = \mathscr{E}_0 \sin \left(\omega t - \frac{2\pi}{3} \right), \qquad (15.1)$$
$$\mathscr{E}_3 = \mathscr{E}_0 \sin \left(\omega t - \frac{4\pi}{3} \right).$$

Обмотку каждой из этих рамок можно замкнуть на свое

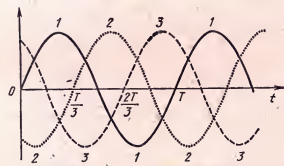


Рис. 15.1. Графики э. д. с., сдвинутых по фазе на 120°.

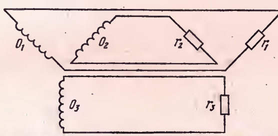


Рис. 15.2. Три независимые линии соединения генератора с потребителями.

нагрузочное сопротивление (рис. 15.2) и получить три синусоидальных переменных тока, между которыми будет строго постоянная разность фаз. Такие три согласованных переменных тока называют трехфазным током. Так как для явлений в электрических цепях важны только разности потенциалов, то можно объединить в один провод по одному

проводу из каждой цепи. В результате получается соединение генератора с потребителем с помощью четырех, а не шести проводов, называемое соединением звездой (рис. 15.3). Если же объединять по одному проводу из каждой цепи на рис. 15.2 попарно, то в результате получается схема соединения генератора с потребителем тремя проводами, называемая соединением треугольником (рис. 15.4).

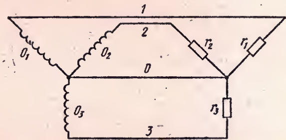


Рис. 15.3. Соединение генератора с потребителями звездой.

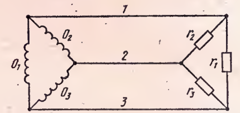


Рис. 15.4. Соединение генератора с потребителями треугольником.

В цепях трехфазного тока напряжения между концами каждой обмотки генератора называются фазными напряжениями, а токи в этих обмотках — фазными токами. Так же называют напряжения и токи в нагрузочных сопротивлениях. Напряжения между проводами *1*, *2*, *3* на рис. 15.3 и между любой парой проводов на рис. 15.4 называются линейными напряжениями, а токи в этих проводах — линейными токами. Легко видеть, что при соединении звездой фазные токи совпадают с линейными токами, а фазные и линейные напряжения различаются. При соединении треугольником, наоборот, совпадают фазные и линейные напряжения, а различаются фазные и линейные токи.

Рассмотрим подробнее соединение звездой. Построим векторные диаграммы токов и напряжений. Предположим, что генератор разомкнут. Тогда фазные напряжения совпадают с соответствующими э. д. с., и поскольку последние сдвинуты по фазе на 120° и 240°, то диаграмма фазных напряжений U_1, U_2 и U_3 имеет вид, показанный на рис. 15.5. Как видно из рис. 15.3, мгновенное значение линейного напряжения между, например, проводами *1* и *2* равно разности мгновенных значений соответствующих фазных напряжений. Поэтому вектор U_{12}, изображающий это напряжение, равен разности векторов U_1 и U_2, изображающих фазные напряжения в первой и второй обмотках (рис. 15.5). Разумеется, вектор U_{12} можно перенести параллельно самому

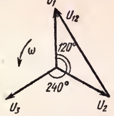


Рис. 15.5. Векторная диаграмма фазных напряжений при соединении звездой.

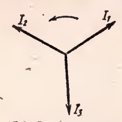


Рис. 15.6. Векторная диаграмма токов в соединении звездой при одинаковых нагрузках.

себе так, чтобы его начало совпало с общим центром вращения векторов. Из рисунка сразу видно, что амплитуда линейного напряжения при соединении звездой в $\sqrt{3}$ раз больше амплитуды фазного. То же самое относится и к действующим значениям этих напряжений. Если, например, фазное напряжение в сети 220 В, то линейное напряжение в этой же сети 380 В.

Так же просто строится векторная диаграмма токов. При одинаковых нагрузках амплитуды токов в проводах *1*, *2*, *3* равны, а токи I_1, I_2, I_3 сдвинуты по фазе на 120° и 240° (рис. 15.6). В нулевом проводе в любой момент ток равен сумме токов I_1, I_2 и I_3 и при симметричной нагрузке, как видно из рис. 15.6, обращается в нуль. В этом случае нулевой провод можно убрать, не изменяя токов в цепи. Если нагрузка несимметрична, то длины векторов, изображающих токи I_1, I_2 и I_3, будут неодинаковы. Теперь в нулевом

проводе будет ток I, амплитуду и фазу которого легко найти с помощью векторной диаграммы, построение которой ясно из рис. 15.7.

Совершенно аналогично может быть построена векторная диаграмма фазных токов в соединении треугольником. При симметричной нагрузке с помощью векторной диаграммы можно убедиться, что амплитуды линейных токов будут в $\sqrt{3}$ раз больше, чем амплитуды фазных токов.

Все приведенные выше результаты можно получить и аналитически, не используя векторных диаграмм. Для этого нужно воспользоваться формулами (15.1) и соответствующими формулами для токов.

В рассмотренных нами схемах и обмотки генератора, и нагрузки соединены одинаково — либо звездой, либо треугольником. Разумеется, можно употреблять и комбинированные схемы, соединяя обмотки генератора звездой, а

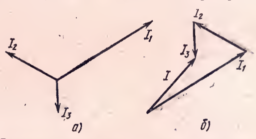


Рис. 15.7. Векторная диаграмма токов при несимметричной нагрузке (*а*) и нахождение вектора I для тока в нулевом проводе (*б*).

нагрузки — треугольником или наоборот. В технике используются различные типы соединений в цепях трехфазного тока, но во всех случаях предпочтительной является симметричная нагрузка фаз, при которой потери будут наименьшими.

Преимущество использования трехфазного тока в технике по сравнению с однофазным заключается в экономии числа проводов и идущего на их изготовление материала. Но самой замечательной особенностью трехфазного тока является то, что он позволяет очень просто создать вращающееся магнитное поле. А с помощью такого поля можно создать простые по конструкции электродвигатели, принцип работы которых заключается в следующем.

Будем вращать подковообразный постоянный магнит так, как показано на рис. 15.8. Вместе с магнитом будет вращаться и создаваемое им магнитное поле. Если в такое поле поместить магнитную стрелку, то она, стремясь установиться вдоль линий индукции магнитного поля, придет во вращение в ту же сторону, в которую вращается поле. Так же будет вести себя и замкнутый виток провода (рис. 15.8).

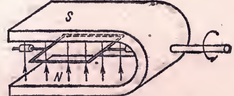


Рис. 15.8. Модель асинхронного двигателя.

Вследствие изменения пронизывающего виток магнитного потока при вращении магнитного поля в витке возникает э. д. с. индукции и индукционный ток. На этот ток со стороны магнитного поля будет действовать сила Ампера. По правилу Ленца индукционный ток в витке направлен так, что взаимодействие этого тока с магнитным полем стремится уменьшить изменение магнитного потока вследствие вращения магнитного поля. Поэтому рамка будет вращаться вслед за магнитным полем. В этом, разумеется, можно убедиться и иначе, если с помощью правила правой руки определить направление индукционного тока в рамке, а затем с помощью правила левой руки определить направление сил Ампера, действующих на отдельные стороны рамки.

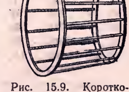


Рис. 15.9. Короткозамкнутый ротор асинхронного двигателя.

Вместо рамки можно взять массивный металлический цилиндр или ротор в виде «беличьего колеса» (рис. 15.9), эквивалентного большому числу соединенных между собой проводящих рамок. При вращении магнитного поля в толще металла цилиндра также будут наводиться замкнутые индукционные токи (вихревые токи, или токи Фуко). Согласно правилу Ленца взаимодействие этих токов с

магнитным полем будет приводить к уменьшению относительной скорости вращения поля и цилиндра.

Выясним, чем отличается поведение во вращающемся магнитном поле магнитной стрелки и короткозамкнутой металлической рамки. При равномерном вращении магнитной стрелки суммарный момент действующих на нее сил должен равняться нулю. Момент сил, действующих на стрелку со стороны магнитного поля, зависит от угла, образованного стрелкой с вектором индукции поля. Этот момент максимален, когда стрелка перпендикулярна полю, и обращается в нуль, когда стрелка направлена по полю. Если на равномерно вращающуюся стрелку никакие другие моменты сил не действуют, то должен быть равен нулю и момент сил, действующих на нее со стороны вращающегося магнитного поля. Следовательно, в любой момент стрелка направлена вдоль поля и вращается синхронно с ним. Если же на стрелку действует тормозящий внешний момент, то стрелка, вращаясь синхронно с полем, будет несколько отставать от него по фазе, так чтобы тормозящий момент уравновешивался моментом сил со стороны магнитного поля. Разумеется, вместо магнитной стрелки можно взять закрепленный на оси постоянный магнит или электромагнит, питаемый постоянным током. Они также будут вращаться синхронно с внешним вращающимся магнитным полем.

Несколько иначе обстоит дело в случае короткозамкнутой рамки или сплошного цилиндра. Индукционный ток зависит от относительной скорости вращения магнитного поля и ротора. При синхронном вращении индукционный ток отсутствует и, следовательно, равен нулю момент сил, действующих на ротор со стороны магнитного поля. Поэтому ротор может вращаться синхронно с полем только тогда, когда никакие тормозящие моменты на него не действуют. При наличии тормозящего момента при равномерном вращении он должен уравновешиваться моментом сил, действующих на индукционные токи в роторе со стороны магнитного поля. Для возникновения этих индукционных токов ротор должен вращаться медленнее магнитного поля. Таким образом, угловая скорость ротора меньше угловой скорости вращения магнитного поля и зависит от величины тормозящего момента. Чем больше тормозящий момент, тем медленнее вращается ротор.

Магнитная стрелка или электромагнит постоянного тока во вращающемся магнитном поле — это модель синхронного двигателя переменного тока, который находит себе применение в тех случаях, когда необходимо иметь строго постоянное, не зависящее от нагрузки число оборотов. Короткозамкнутый ротор во вращающемся магнитном поле — это модель асинхронного двигателя переменного тока, угловая скорость вращения ротора которого зависит от механической нагрузки. В силу исключительной простоты конструкции и высокой надежности асинхронные двигатели получили широкое распространение в технике.

Опишем теперь способ получения вращающегося магнитного поля в электродвигателях переменного тока.

Предположим, что у нас есть равномерно вращающееся против часовой стрелки в плоскости xy магнитное поле, вектор индукции $\boldsymbol{B}$ которого не меняется по величине (рис. 15.10). Из этого рисунка видно, что такое поле можно рассматривать как результат сложения двух магнитных полей, индукция одного из которых направлена вдоль оси x и меняется со временем по закону

$$B_x(t) = B\cos\omega t, \qquad (15.2)$$

а индукция другого направлена по оси y и имеет вид

$$B_y(t) = B\sin\omega t = B\cos\left(\omega t - \frac{\pi}{2}\right), \qquad (15.3)$$

т. е. отстает по фазе от B_x на $\pi/2$. Такие поля легко получить, расположив взаимно перпендикулярно две одинаковые катушки и пропуская по ним синусоидальные токи одинаковой амплитуды, но сдвинутые по фазе на $\pi/2$. Сумма этих полей и дает вращающееся магнитное поле.

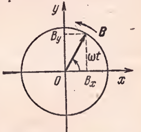


Рис. 15.10. Вращающееся магнитное поле.

Совершенно аналогично можно получить вращающееся магнитное поле с помощью трехфазного переменного тока. Для этого нужно три одинаковые катушки расположить так, чтобы их оси лежали в одной плоскости под углом 120° друг к другу и пересекались в одной точке, и включить катушки в сеть трехфазного тока по схеме звезды или треугольника. Тогда магнитное поле,

создаваемое каждой катушкой, будет направлено вдоль оси соответствующей катушки и будет зависеть от времени в соответствии с формулами (15.1):

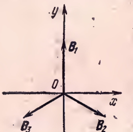


Рис. 15.11. Получение вращающегося магнитного поля при использовании трехфазного тока.

$$B_1(t)=B_{01}\sin\omega t,$$
$$B_2(t)=B_{02}\sin\left(\omega t-\frac{2\pi}{3}\right), \qquad (15.4)$$
$$B_3(t)=B_{03}\sin\left(\omega t-\frac{4\pi}{3}\right).$$

Амплитудные значения этих полей B_{01}, B_{02} и B_{03} равны между собой. Обозначим их через B_0. Результирующее поле

$$B(t)=B_1(t)+B_2(t)+B_3(t) \qquad (15.5)$$

имеет постоянную величину, равную $^3/_2 B_0$, и равномерно вращается в плоскости осей катушек с угловой скоростью ω. Чтобы убедиться в этом, спроектируем вектор индукции результирующего поля B на оси x и y (рис. 15.11):

$$B_x=B_2\cos 30°-B_3\cos 30°=$$
$$=\frac{\sqrt{3}}{2}B_0\left[\sin\left(\omega t-\frac{2\pi}{3}\right)-\sin\left(\omega t-\frac{4\pi}{3}\right)\right].$$

Используя формулу разности синусов двух углов, получим

$$B_x=\frac{3}{2}B_0\cos\omega t. \qquad (15.6)$$

Аналогично,

$$B_y=B_1-B_2\sin 30°-B_3\sin 30°=$$
$$=B_0\left[\sin\omega t-\frac{1}{2}\sin\left(\omega t-\frac{2\pi}{3}\right)-\frac{1}{2}\sin\left(\omega t-\frac{4\pi}{3}\right)\right].$$

Преобразуя второе и третье слагаемые в квадратных скобках с помощью формулы суммы синусов, найдем

$$B_y=\frac{3}{2}B_0\sin\omega t. \qquad (15.7)$$

Сравнивая (15.6) и (15.7) с формулами (15.2) и (15.3), убеждаемся, что поле B действительно равномерно вращается с

угловой скоростью ω. Направление вращения магнитного поля, а следовательно и ротора электродвигателя, можно изменить на противоположное, если поменять местами концы любой пары проводов, присоединенных к катушкам статора, создающим магнитное поле.

Заканчивая изучение принципа действия электродвигателей переменного тока, отметим, что синхронные машины являются обратимыми, т. е. могут быть использованы и как генераторы. Если не подавать напряжение на обмотки статора, а ротор (электромагнит) привести во вращение, то в обмотках статора будет индуцироваться переменное трехфазное напряжение. А асинхронный двигатель подобен трансформатору, у которого вторичная обмотка выполнена подвижной. Что же касается физической сущности явлений, то в обоих случаях она одинакова, поскольку токи и во вторичной обмотке трансформатора, и в роторе асинхронного двигателя имеют чисто индукционное происхождение.

4. КОЛЕБАНИЯ И ВОЛНЫ

СОБСТВЕННЫЕ КОЛЕБАНИЯ В МЕХАНИЧЕСКИХ СИСТЕМАХ И ЭЛЕКТРИЧЕСКИХ ЦЕПЯХ

§ 1. Свободные колебания гармонического осциллятора

Среди разнообразных физических явлений в окружающем нас мире мы часто наблюдаем периодические или почти периодические процессы: восход и заход солнца, волнение на море, колебания маятника часов, переменный электрический ток, электромагнитные волны, колебания молекул в твердом теле,— примеры можно было бы продолжать до бесконечности.

Колебательные явления обладают общими чертами и даже подчиняются одинаковым закономерностям, несмотря на то, что могут иметь совершенно разную физическую природу. Самая характерная черта колебательных движений, отличающая их от других явлений, состоит в том, что колебательные движения многократно повторяются или приблизительно повторяются через определенные промежутки времени. Универсальность законов колебательных процессов позволяет с единой точки зрения рассматривать различные по физической природе колебания, встречающиеся в разнообразных физических явлениях и технических устройствах.

Единый подход к изучению колебаний разной физической природы позволяет глубже проанализировать любое конкретное явление, выявить аналогию между совершенно разными по своей природе явлениями, найти общий язык для их описания и в конечном счете почувствовать единство физического мира.

Со времен Ньютона развитие физики происходило таким образом, что при изучении любого нового явления — электрического, оптического — для него прежде всего пытались

придумать механическую аналогию, т. е. объяснить его с точки зрения законов механики. Например, Кельвин говорил, что понимает явление, если может составить для него механическую модель. Максвелл приложил много усилий для того, чтобы объяснить с помощью механических представлений найденные им уравнения электромагнитного поля. Однако многие современные физики и инженеры уже предпочли бы сказать, что понимают механическое явление, если создали для него электрическую модель. Именно при изучении колебательных процессов пришло в физику отчетливое понимание того, что явления разной природы, несмотря на внешнее сходство, не сводимы друг к другу, однако могут подчиняться одинаковым законам и описываться одними и теми же уравнениями.

Любая система, способная совершать колебательное движение, описывается некоторой физической величиной, отклонение которой от равновесного значения зависит от времени по периодическому или почти периодическому закону. Определение периодической функции таково: функция $f(t)$ называется периодической с периодом T, если $f(t+T)=f(t)$ при любом значении t. В случае механических колебательных процессов, например колебаний груза, подвешенного на пружине, такой величиной является смещение груза из положения равновесия. В случае электрических систем, например колебательного контура, такой величиной является ток в катушке или заряд на обкладках конденсатора.

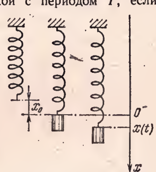


Рис. 1.1. Положение равновесия и колебания груза на пружине.

Покажем, что колебания груза, подвешенного на пружине, и колебания в контуре, состоящем из катушки индуктивности и конденсатора, описываются одинаковыми уравнениями. Пусть на пружине жесткости k подвешено тело массы m (рис. 1.1). Рассмотрим вертикальное движение тела, которое будет происходить под действием силы упругости пружины и силы тяжести после толчка. Массу пружины предполагаем настолько малой, чтобы ею можно

было пренебречь при колебаниях. Поместим начало отсчета по оси x в точку, соответствующую равновесному положению тела. В этом положении благодаря действию силы тяжести пружина уже растянута на величину x_0, определяемую соотношением

$$mg=kx_0. \tag{1.1}$$

Поэтому при смещении груза на величину x из положения равновесия действующая на тело со стороны пружины сила равна $-k(x+x_0)$. Обозначая ускорение тела a, равное второй производной смещения x по времени, через $\ddot{x}$, запишем второй закон Ньютона в виде

$$m\ddot{x}=-k(x+x_0)+mg. \tag{1.2}$$

С учетом (1.1) это уравнение переписывается в виде

$$m\ddot{x}=-kx. \tag{1.3}$$

Вводя обозначение

$$\omega_0^2=\frac{k}{m}, \tag{1.4}$$

представим уравнение движения тела (1.3) в виде

$$\ddot{x}+\omega_0^2x=0. \tag{1.5}$$

Перейдем к рассмотрению колебаний в контуре, содержащем последовательно соединенные катушку индуктивности L и конденсатор емкости C (рис. 1.2). При обходе контура (например, против часовой стрелки) сумма напряжений на катушке индуктивности U_L и конденсаторе U_C в такой последовательной цепи равна нулю:

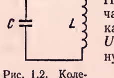


Рис. 1.2. Колебательный контур.

$$U_L+U_C=0. \tag{1.6}$$

Напряжение на конденсаторе U_C связано с зарядом пластины q и емкостью C соотношением $U_C=q/C$. Напряжение на индуктивности U_L в любой момент времени равно по величине и противоположно по знаку э. д. с. самоиндукции, поэтому $U_L=L\frac{dI}{dt}$. Ток в цепи I, как видно из рис. 1.2, равен скорости изменения заряда пластины конденсато-

ра: $I = dq/dt$. Подставляя ток в выражение для напряжения на катушке и обозначая вторую производную заряда конденсатора q по времени через $\ddot{q}$, получим $U_L = L\ddot{q}$. Теперь выражение (1.6) принимает вид

$$L\ddot{q} + \frac{q}{C} = 0. \tag{1.7}$$

Вводя обозначение

$$\omega_0^2 = \frac{1}{LC}, \tag{1.8}$$

перепишем уравнение (1.7) в виде

$$\ddot{q} + \omega_0^2 q = 0. \tag{1.9}$$

Видно, что уравнение (1.5), описывающее колебания груза на пружине, и уравнение (1.9), описывающее электромагнитные колебания в контуре, совпадают.

К такому же точно уравнению мы придем, рассматривая малые колебания математического маятника, физического маятника, т. е. любого тела, которое может вращаться вокруг горизонтальной оси под действием силы тяжести, крутильные колебания диска или коромысла, подвешенного на упругой нити, и т. д.

Колебания любой физической природы, описываемые таким уравнением, называются гармоническими, а совершающая такие колебания система — гармоническим осциллятором.

Решение уравнения гармонических колебаний имеет вид

$$f(t) = A\cos(\omega_0 t + \alpha), \tag{1.10}$$

где A и α — произвольные постоянные. Функция $f(t)$ (1.10) является решением дифференциального уравнения (1.5) или (1.9) при любых значениях A и α. При колебаниях груза на пружине $f(t)$ — это смещение груза из положения равновесия, т. е. $f(t) = x(t)$, а при электромагнитных колебаниях в контуре $f(t)$ есть заряд на обкладке конденсатора, т. е. $f(t) = q(t)$.

Циклическая частота гармонических колебаний ω_0 не зависит от амплитуды A и определяется параметрами изучаемой системы: при механических колебаниях груза на пружине значение ω_0 дается формулой (1.4), а при электро-

магнитных колебаниях в контуре — формулой (1.8). Величина постоянных A и α в выражении (1.10), имеющих смысл амплитуды и начальной фазы колебаний, зависит от начального состояния системы. Например, при механических колебаниях значения A и α определяются заданием начального механического состояния системы, т. е. значений смещения груза X_0 и его скорости V_0 при $t=0$. С помощью (1.10) при $t=0$ получаем

$$X_0 = A\cos\alpha,$$
$$V_0 = -A\omega_0 \sin\alpha. \qquad (1.11)$$

Решая эту систему уравнений относительно A и α, получаем

$$A = \sqrt{X_0^2 + \frac{V_0^2}{\omega_0^2}}, \quad \operatorname{tg}\alpha = -\frac{V_0}{\omega_0 X_0}. \qquad (1.12)$$

Для определения постоянных A и α в случае электромагнитных колебаний в контуре нужно задать значения заряда на конденсаторе q_0 и тока в цепи I_0 при $t=0$.

Убедиться в том, что функция (1.10) является решением уравнения гармонических колебаний (1.5) или (1.9), можно непосредственной подстановкой. Но можно это сделать и иначе, воспользовавшись удобным графическим методом изображения колебаний. Рассмотрим равномерное движение точки по окружности радиуса A с угловой скоростью

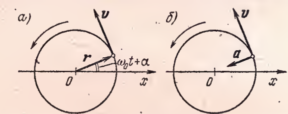


Рис. 1.3. Векторы скорости и ускорения при равномерном движении точки по окружности.

ω_0 (рис. 1.3, *а*). Пусть в начальный момент радиус-вектор этой точки образует угол α с осью x. Спроектируем теперь на эту ось радиус-вектор движущейся точки $\boldsymbol{r}$, ее скорость $\boldsymbol{v}$ и ускорение $\boldsymbol{a}$. Учитывая, что при равномерном движении точки по окружности ее скорость направлена по касательной, а ускорение — к центру окружности (рис. 1.3, *б*),

получим

$$x(t)=A\cos(\omega_0 t+\alpha),$$
$$v_x(t)=-\omega_0 A\sin(\omega_0 t+\alpha), \qquad (1.13)$$
$$a_x(t)=-\omega_0^2 A\cos(\omega_0 t+\alpha).$$

Мы воспользовались тем, что при движении по окружности величина скорости v связана с радиусом окружности A и угловой скоростью ω_0 соотношением $v=\omega_0 A$, а величина ускорения — соотношением $a=\omega_0^2 A$. Из формул (1.13) видно, что проекция ускорения $a_x(t)$ в любой момент времени пропорциональна смещению $x(t)$, точно так же, как и в уравнении (1.5):

$$a_x(t)=-\omega_0^2 x(t). \qquad (1.14)$$

Отсюда следует, что уравнение (1.5) описывает движение, происходящее по синусоидальному закону (1.13).

Подчеркнем, что при гармонических колебаниях любой физической природы, которые происходят по закону (1.10), частота ω_0 оказывается зависящей только от свойств системы, в которой происходят колебания, но не зависит от амплитуды колебаний. В одной и той же системе могут происходить колебания определенной частоты, которая, например, дается формулами (1.4) или (1.8), но разной амплитуды. Амплитуда колебаний A и начальная фаза α определяются не свойствами самой системы, а тем способом, каким в системе вызваны колебания. Колебания, происходящие в системе в результате вывода ее из состояния равновесия, после чего система предоставляется самой себе, будем называть собственными колебаниями. В отсутствие трения собственные колебания иногда называют свободными.

Рассмотрим энергетические превращения, происходящие при свободных гармонических колебаниях.

При механических колебаниях груза на пружине происходит периодическое превращение кинетической энергии движущегося груза $E_к$ и потенциальной энергии $E_п$ системы, которая состоит из потенциальной энергии деформированной пружины и потенциальной энергии груза в поле тяжести. Потенциальная энергия деформированной пружины пропорциональна квадрату ее удлинения $x+x_0$ (рис. 1.1) и, следовательно, равна $\frac{1}{2}k(x+x_0)^2$. Потенциальная энер-

гия груза в поле тяжести равна $-mgx+C$. Выберем для удобства произвольную постоянную C таким образом, чтобы полная потенциальная энергия системы была равна нулю в положении равновесия:

$$\frac{1}{2}kx_0^2+C=0,$$

откуда $C=-\frac{1}{2}kx_0^2$, и потенциальная энергия системы $E_п$ в произвольной точке x выражается формулой

$$E_п=\frac{1}{2}k(x+x_0)^2-mgx-\frac{1}{2}kx_0^2=\frac{1}{2}kx^2. \qquad (1.15)$$

Полная механическая энергия системы $E=E_к+E_п$ при колебаниях остается неизменной, так как система консервативна. В этом можно убедиться и непосредственно, подставляя смещение x и скорость v из (1.13) в выражение для энергии:

$$E=\frac{1}{2}mv_x^2+\frac{1}{2}kx^2=\frac{1}{2}m\omega_0^2A^2\sin^2(\omega_0t+\alpha)+$$
$$+\frac{1}{2}kA^2\cos^2(\omega_0t+\alpha)=\frac{1}{2}kA^2=\frac{1}{2}m\omega_0^2A^2. \qquad (1.16)$$

Из этой формулы видно, что неизменная полная энергия системы E совпадает с потенциальной энергией $E_п$ в точках наибольшего отклонения от положения равновесия, т. е. при $x=\pm A$, и совпадает с кинетической энергией $E_к$ при прохождении груза через положение равновесия, где его скорость v_x равна $\pm\omega_0 A$. При взаимных превращениях потенциальная и кинетическая энергии совершают гармонические колебания с одинаковой амплитудой $E/2$ в противофазе друг с другом и с частотой $2\omega_0$. Чтобы убедиться в этом, преобразуем выражения для кинетической и потенциальной энергий с помощью формул для тригонометрических функций половинного аргумента:

$$E_к(t)=\frac{1}{2}m\omega_0^2A^2\sin^2(\omega_0t+\alpha)=\frac{E}{2}[1-\cos 2(\omega_0t+\alpha)],$$
$$(1.17)$$
$$E_п(t)=\frac{1}{2}kA^2\cos^2(\omega_0t+\alpha)=\frac{E}{2}[1+\cos 2(\omega_0t+\alpha)].$$

На рис. 1.4 приведены графики зависимости от времени смещения груза $x(t)$, кинетической $E_к(t)$ и потенциальной $E_п(t)$ энергий.

В случае электромагнитных колебаний в контуре происходят взаимные превращения энергии электрического

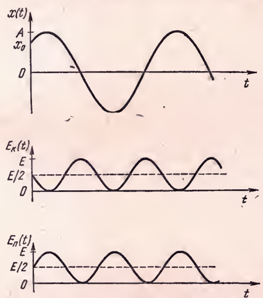


Рис. 1.4. Графики смещения, кинетической и потенциальной энергий при колебаниях.

поля в конденсаторе $W_C=\frac{1}{2}q^2/C$ и энергии магнитного поля в катушке индуктивности $W_L=\frac{1}{2}LI^2$. Полная энергия колебаний W, равная W_C+W_L, остается неизменной. Ее можно выразить через амплитуду колебаний заряда конденсатора Q:

$$W=\frac{1}{2C}Q^2. \quad (1.18)$$

Наглядное представление о процессе колебаний можно получить с помощью так называемых фазовых диаграмм. Механическое состояние совершающего колебания тела

определяется заданием его координаты x и скорости v_x. Если на плоскости ввести систему координат и отложить x по оси абсцисс и v_x по оси ординат, то состояние системы будет изображаться точкой на этой плоскости. При изменении механического состояния изображающая его точка будет в этой плоскости двигаться по некоторой линии. Если рассматриваемая система возвращается в исходное состояние, то соответствующая такому движению линия замыкается. Плоскость x, v_x называется фазовой плоскостью, а кривая, по которой движется изображающая точка при изменении механического состояния системы, называется фазовой траекторией.

Построим фазовые траектории для гармонического осциллятора. Поскольку при свободных колебаниях энергия системы сохраняется, то все точки замкнутой фазовой траектории соответствуют одному и тому же значению энергии. Поэтому уравнение фазовой траектории представляет собой уравнение закона сохранения энергии:

$$\frac{1}{2}kx^2+\frac{1}{2}mv_x^2=E. \qquad (1.19)$$

Переписывая это уравнение в виде

$$\frac{x^2}{2E/k}+\frac{v_x^2}{2E/m}=1, \qquad (1.20)$$

убеждаемся, что фазовая траектория гармонического осциллятора представляет собой эллипс с полуосями $\sqrt{2E/k}$ и $\sqrt{2E/m}$ (рис. 1.5). При колебаниях состояние осциллятора меняется таким образом, что изображающая точка движется по эллипсу по часовой стрелке и совершает полный оборот за время, равное периоду колебаний $T=2\pi/\omega_0$. В этом легко убедиться с помощью формул (1.13), дающих зависимость x и v_x от времени. Из этих формул, разумеется, можно получить и само уравнение фазовой траектории (1.20), если исключить из них время. Для этого нужно обе части первой из формул (1.13) разделить на A, второй — на $\omega_0 A$, возвести получившиеся выражения в квадрат и сложить.

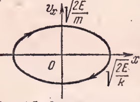


Рис. 1.5. Фазовая траектория гармонического осциллятора.

В результате получим

$$\frac{x^2}{A^2}+\frac{v_x^2}{(\omega_0 A)^2}=1,$$

что совпадает с уравнением (1.20), ибо $E=\frac{1}{2}kA^2=\frac{1}{2}m\omega_0^2A^2$.

Интересно сопоставить вид фазовой траектории с графиком потенциальной энергии (рис. 1.6). На верхней части рисунка изображена потенциальная энергия осциллятора и показаны два значения полной энергии системы E_1 и E_2. На нижней части изображены две фазовые траектории осциллятора, соответствующие колебаниям с такими значениями энергии. Скорость обращается в нуль в тех точках, где потенциальная энергия становится равной полной энергии, т. е. в точках максимального смещения из положения равновесия. Величина скорости максимальна при прохождении положения равновесия $x=0$, где потенциальная энергия обращается в нуль. Масштаб графика фазовой траектории по оси v_x произволен и не связан с графиком потенциальной энергии. Удобно масштаб графика выбрать так, чтобы одинаковые по величине отрезки соответствовали единице по оси x и ω_0 по оси v_x. Тогда при любой амплитуде колебаний A полуоси эллипса на фазовой диаграмме A и $\omega_0 A$ будут одинаковы и эллипс превратится в окружность (рис. 1.7). Точка, изображающая состояние осциллятора, движется по этой окружности по часовой стрелке с постоянной скоростью. Из рис. 1.7 видна связь движения изображающей точки в фазовой плоскости с временно́й зависимостью координаты $x(t)$ и скорости $v_x(t)$ осциллятора. При построении фазовых диаграмм мы будем выбирать масштаб по осям именно таким образом.

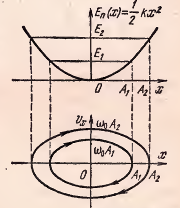


Рис. 1.6. Потенциальная энергия и фазовая траектория.

В дальнейшем мы увидим, что метод фазовых траекторий оказывается очень эффективным и при изучении более

сложных, чем гармонические колебания, движений. Он дает наглядное представление о характере движения даже

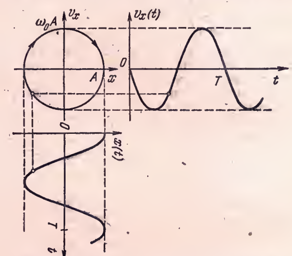


Рис. 1.7. Связь фазовой траектории с графиками смещения и скорости.

тогда, когда не удается получить аналитическое решение уравнений движения.

§ 2. Затухающие колебания

Свободные колебания, рассмотренные в предыдущем параграфе, представляют собой некоторую идеализацию. В реальных системах механическое движение всегда происходит в какой-либо внешней среде, которая оказывает сопротивление движению. Наличие сил трения приводит к рассеянию, или, как говорят, к диссипации, механической энергии. Диссипация энергии колебаний происходит в любых реальных колебательных системах. Например, в колебательном контуре всегда имеется активное сопротивление, на котором происходит выделение тепла при прохождении тока. Поэтому собственные колебания фактически всегда являются затухающими.

Рассмотрим сначала затухающие механические колебания. При движении тела в среде при малых скоростях

силу сопротивления можно считать пропорциональной скорости тела:

$$\boldsymbol{F}_{\text{тр}} = -\beta \boldsymbol{v}. \tag{2.1}$$

Поэтому уравнение (1.3) колебания груза, подвешенного на пружине, при наличии трения будет иметь вид

$$m\ddot{x} = -kx - \beta\dot{x}, \tag{2.2}$$

где через $\dot{x}$ обозначена производная смещения x по времени, т. е. проекция скорости тела. Вводя обозначения

$$\omega_0^2 = \frac{k}{m}, \quad 2\gamma = \frac{\beta}{m}, \tag{2.3}$$

перепишем уравнение (2.2) в виде

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = 0. \tag{2.4}$$

Рис. 2.1. Колебательный контур с сопротивлением.

Покажем, что точно таким же уравнением описываются колебания в контуре, содержащем последовательно соединенные конденсатор емкости C, катушку индуктивности L и резистор сопротивления R (рис. 2.1). Сумма напряжений на отдельных элементах цепи U_C, U_L и U_R в любой момент времени равна нулю:

$$U_L + U_R + U_C = 0. \tag{2.5}$$

Подставляя в (2.5) $U_C = q/C$, $U_L = L\ddot{q}$ и $U_R = R\dot{q}$, получим

$$L\ddot{q} + R\dot{q} + \frac{q}{C} = 0. \tag{2.6}$$

Вводя обозначения

$$\omega_0^2 = \frac{1}{LC}, \quad 2\gamma = \frac{R}{L}, \tag{2.7}$$

перепишем уравнение (2.6) в виде

$$\ddot{q} + 2\gamma\dot{q} + \omega_0^2 q = 0, \tag{2.8}$$

что совпадает с уравнением (2.4). Таким образом, и при наличии затухания как механические колебания груза на пружине, так и колебания в контуре происходят по одинаковому закону.

Не будем пока решать уравнение (2.4) или (2.8), а попробуем выяснить, как влияет наличие сопротивления на колебательное движение. Будем при этом считать, что затухание мало настолько, что связанная с ним потеря энергии системы за период колебания мала по сравнению с энергией колебаний. Для определенности будем говорить о механических колебаниях. Согласно закону сохранения энергии изменение механической энергии системы, совершающей колебания, равно работе силы трения:

$$\Delta E = (\boldsymbol{F}_{\text{тр}} \cdot \Delta \boldsymbol{r}).$$

Подставляя сюда силу трения из (2.1) и учитывая, что $\Delta x = v_x \Delta t$, получим

$$\Delta E = -\beta v_x \Delta x = -\beta v_x^2 \Delta t. \tag{2.9}$$

Из соотношения (2.9) в пределе при $\Delta t \to 0$ видно, что скорость изменения энергии колебаний dE/dt пропорциональна квадрату скорости и поэтому может быть выражена через кинетическую энергию $E_{\text{к}} = \frac{1}{2} m v_x^2$:

$$\frac{dE}{dt} = -\beta v_x^2 = -\frac{2\beta}{m} \frac{m v_x^2}{2} = -\frac{2\beta}{m} E_{\text{к}}. \tag{2.10}$$

Из формулы (2.10) видно, что диссипация энергии в течение периода колебаний происходит неравномерно, так как кинетическая энергия $E_{\text{к}}$ осциллирует. Нас интересует потеря энергии колебаний за период. Ее можно охарактеризовать средним значением $\langle dE/dt \rangle$. Усредняя выражение (2.10) по периоду колебаний, можем написать

$$\left\langle \frac{dE}{dt} \right\rangle = -\frac{2\beta}{m} \langle E_{\text{к}} \rangle. \tag{2.11}$$

Ввиду малости затухания при нахождении среднего значения кинетической энергии $\langle E_{\text{к}} \rangle$ можно воспользоваться формулой (1.17), справедливой для свободных колебаний. Так как среднее значение $\cos 2(\omega_0 t + \alpha)$ за период равно нулю, то $\langle E_{\text{к}} \rangle = E/2$.

Подставляя это значение в формулу (2.11) и используя обозначение (2.3), получим

$$\left\langle \frac{dE}{dt} \right\rangle = -2\gamma E. \tag{2.12}$$

Формула (2.12) показывает, что усредненная по периоду колебаний скорость изменения энергии $\langle dE/dt \rangle$, характеризующая «сглаженное» поведение энергии колебаний, когда нас не интересуют детали ее изменения на протяжении одного периода колебаний, пропорциональна самой энергии E. Поэтому для промежутков времени, больших по сравнению с периодом, знак $\langle \ldots \rangle$ усреднения по периоду в (2.12) можно опустить. В результате получается уравнение, решение которого показывает, что изменение $E(t)$ происходит по экспоненциальному закону:

$$E(t) = E_0 e^{-2\gamma t}. \tag{2.13}$$

Здесь E_0 — значение энергии системы в начальный момент. Но энергия колебаний пропорциональна квадрату амплитуды. Поэтому изменение амплитуды колебаний за промежутки времени, большие по сравнению с периодом, дается выражением

$$A = A_0 e^{-\gamma t}, \tag{2.14}$$

где A_0 — начальная амплитуда колебаний. Зависимость амплитуды от времени показана пунктиром на рис. 2.2.

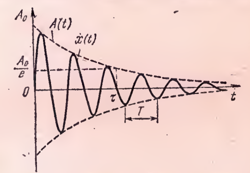


Рис. 2.2. График затухающих колебаний.

Как видно из (2.14), амплитуда убывает в e раз за время τ, равное $1/\gamma$, независимо от начального значения амплитуды. Это время τ носит название времени жизни колебаний, хотя, как видно из формулы (2.14), колебания продолжаются бесконечно долго. Использованное нами предположение о малости затухания означает, что время жизни

колебаний τ велико по сравнению с периодом T: $\tau \gg T$. Другими словами, за время τ происходит большое число колебаний. Отметим, что в данном случае движение, строго говоря, не является периодическим. Под периодом колебаний T здесь понимают промежуток времени между двумя последовательными максимальными отклонениями от равновесия. Фазовая траектория затухающего колебания при наличии трения, пропорционального скорости, приведена на рис. 2.3. Она представляет собой незамкнутую кривую — спираль, закручивающуюся вокруг начала координат. При малом затухании, когда осциллятор за время жизни τ успевает совершить большое число колебаний, такое же число витков накручивает спираль на фазовой плоскости.

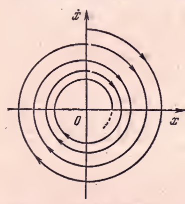


Рис. 2.3. Фазовая траектория осциллятора с трением, пропорциональным скорости.

Затухание колебаний влияет и на период, приводя к его возрастанию по сравнению с периодом свободных колебаний в той же системе. Однако при малом затухании увеличение периода колебаний очень мало. Но при сильном затухании колебаний вообще может не быть: выведенная из равновесия система вследствие большого трения будет апериодически, т. е. без осцилляций, приближаться к положению равновесия.

Уравнение затухающих колебаний (2.4) имеет точное решение. Непосредственной подстановкой можно убедиться, что оно имеет вид

$$x(t) = A_0 e^{-\gamma t} \cos(\omega_1 t + \alpha), \quad \omega_1 = \sqrt{\omega_0^2 - \gamma^2}, \qquad (2.15)$$

где A_0 и α — произвольные постоянные, значение которых определяется из начальных условий. При малом затухании, когда $\gamma \ll \omega_0$, частота ω_1 практически совпадает с частотой свободных колебаний $\omega_0 = \sqrt{k/m}$, а стоящий перед

косинусом множитель $A_0 e^{-\gamma t}$ можно рассматривать как медленно меняющуюся со временем амплитуду колебаний.

Совершенно аналогично будет происходить затухание колебаний в контуре с сопротивлением, которые описываются уравнением (2.8).

Экспоненциальный характер затухания колебаний связан с тем, что вызывающая это затухание сила трения пропорциональна скорости. При другой зависимости силы трения от скорости закон затухания колебаний будет иным. Рассмотрим случай сухого трения, когда от скорости зависит только направление силы трения, а ее величина практически постоянна. Пусть на горизонтальный стержень насажен просверленный по диаметру шар массы m, прикрепленный к пружине жесткости k (рис. 2.4). Сила трения, равная μmg, направлена в сторону, противоположную скорости $\dot{x}$. Поэтому уравнение движения шара записывается следующим образом:

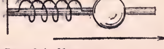


Рис. 2.4. Между шаром и стержнем — сухое трение.

$$\begin{aligned} m\ddot{x} &= -kx - \mu mg \quad \text{при} \quad \dot{x} > 0, \\ m\ddot{x} &= -kx + \mu mg \quad \text{при} \quad \dot{x} < 0. \end{aligned} \tag{2.16}$$

Таким образом, для нахождения движения шара необходимо решать два уравнения, которые сменяют друг друга, когда меняется направление движения шара, т. е. знак скорости $\dot{x}$. Пусть в начальный момент шар смещен из положения равновесия влево на некоторую величину A, а скорость его равна нулю. Если при этом упругая сила пружины меньше, чем максимально возможное значение силы трения покоя μmg, то шар будет оставаться в покое и дальше. Таким образом, вблизи положения равновесия $x=0$, соответствующего ненапряженной пружине, существует область «застоя» шириной $2\mu mg/k$, в любой точке которой шар может находиться в покое. Если же начальное смещение сдвинутого влево шара A больше, чем $\mu mg/k$, то отпущенный шар начнет двигаться направо, и его движение будет определяться первым из уравнений (2.16). Это уравнение описывает гармонические колебания с частотой $\omega_0 = \sqrt{k/m}$. Наличие постоянной силы в правой части этого уравнения,

не меняя частоты колебаний, приводит к сдвигу положения равновесия (вспомним, что в разобранном выше примере колебаний груза на пружине в поле тяжести колебания происходят с той же частотой, что и в невесомости, но около нового положения равновесия).

Записывая первое из уравнений (2.16) в виде

$$m\ddot{x} = -k(x - x_0), \tag{2.17}$$

находим, что сдвиг положения равновесия x_0, относительно которого происходят описываемые этим уравнением колебания, равен

$$x_0 = -\frac{\mu m g}{k}.$$

Так как $x_0 < 0$, то положение равновесия смещено влево. После того как шар дойдет до крайнего правого положения

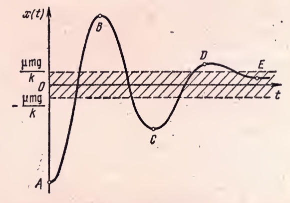


Рис. 2.5. График затухающих колебаний при сухом трении. Область застоя заштрихована.

и его скорость обратится в нуль, он начнет двигаться налево, $\dot{x}$ станет меньше нуля и движение шара будет определяться вторым из уравнений (2.16). Это уравнение в свою очередь описывает гармонические колебания с той же частотой ω_0, происходящие около другого положения равновесия, сдвинутого относительно точки $x=0$ на ту же величину в противоположную сторону. После того как шар придет в крайнее левое положение, дальнейшее его

движение снова будет описываться первым из уравнений (2.16), и т. д. Зависимость смещения шара от времени $x(t)$ показана на рис. 2.5. Сначала шар идет слева направо, чему соответствует отрезок синусоиды AB, изображающий незатухающие колебания около положения $-\mu mg/k$. Движение справа налево изображается отрезком синусоиды BC, соответствующим колебанию около положения $\mu mg/k$, и т. д. В результате чередования кусков синусоид, описывающих незатухающие колебания около двух чередующихся положений равновесия, получается кривая, описывающая затухающее движение. Очевидно, что рано или поздно скорость шара обратится в нуль в тот момент, когда он будет находиться внутри области застоя (точка E на рис. 2.5), и на этом его движение прекратится. В отличие от затухающих колебаний при сопротивлении, пропорциональном скорости, здесь колебания полностью прекращаются через конечное время.

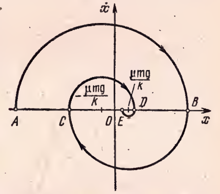


Рис. 2.6. Фазовая диаграмма колебаний при сухом трении, соответствующая графику на рис. 2.5.

Наглядное представление о рассмотренных колебаниях при наличии сухого трения можно получить и с помощью фазовой диаграммы (рис. 2.6). Начальное состояние изображается точкой A на оси x. Движению шара слева направо соответствует часть фазовой траектории AB, представляющая собой половину окружности, центр которой находится на оси x в точке $-\mu mg/k$. Дальнейшему движению справа налево соответствует половина окружности BC, центр которой находится в точке $\mu mg/k$ на оси x, и т. д. Вся фазовая траектория состоит из таких половинок окружностей с чередующимися центрами. Она обрывается в точке E на оси x, как только достигает области застоя.

Рассмотренные особенности затухающих колебаний позволяют понять происхождение погрешностей у стрелочных измерительных приборов, связанных с успокоением их

подвижной системы при измерениях. Неизбежно присутствующее сухое трение приводит к существованию области застоя около положения равновесия стрелки прибора при проведении измерения. Из рис. 2.5 и 2.6 видно, что остановка после колебаний может произойти в любой точке области застоя в зависимости от начальных условий. Поэтому остановка стрелки прибора происходит не точно на том делении шкалы, которое соответствует измеряемой величине, а в какой-либо точке области застоя вблизи этого деления. Для уменьшения ошибки измерений сухое трение стремятся свести к минимуму. Один из способов уменьшения сухого трения — применение смазки. При этом трение становится пропорциональным скорости и затухание происходит в соответствии с законом (2.15). Стрелка при этом должна остановиться в положении равновесия. Чтобы успокоение подвижной системы прибора не происходило слишком долго, применяются так называемые демпфирующие устройства, гасящие колебания. Эти устройства не должны ухудшать точность прибора, поэтому вводимое ими трение должно быть пропорционально скорости.

§ 3. Энергетические превращения при собственных колебаниях

При изучении колебаний, как и любого другого физического явления, мы всегда вынуждены упрощать рассматриваемую систему, стремясь тем не менее сохранить в выбранной идеализированной модели наиболее важные черты явления. Однако никакую идеализацию нельзя продолжать до бесконечности, нужно всегда отдавать себе отчет, до каких пределов остается справедливой выбранная модель. Но и в рамках выбранной модели иногда еще остаются вопросы, связанные с условиями применимости приближений, использованных при конкретных расчетах.

Проанализируем с этой точки зрения те приближения, которые были использованы в § 1 при изучении колебаний груза, подвешенного на упругой пружине. Задумаемся над вопросом, в чем смысл сделанного там предположения о малости массы пружины по сравнению с массой груза. Ведь при выводе уравнения движения груза (1.5) предположение об этом, казалось бы, нигде не использовалось. Действительно, мы воспользовались вторым законом Нью-

тона для груза (1.2), в который входит масса груза m, но не входит масса пружины. Однако уравнение колебаний груза (1.5) все-таки справедливо только в том случае, когда масса пружины достаточно мала.

На первый взгляд может показаться, что дело здесь только в том, что массивная пружина будет растянута еще и под действием собственной тяжести, так что действующая на груз со стороны пружины сила уже не будет равна $k(x+x_0)$, как в уравнении (1.2). Однако дело совсем не в этом. И при горизонтальном расположении пружины (рис. 2.4) в отсутствие трения уравнение (1.5) справедливо лишь тогда, когда масса пружины мала по сравнению с массой груза. В противном случае нужно учитывать движение самой пружины. В самом деле, при получении закона движения (1.5) предполагается, что если конец пружины оттянут на расстояние x, то действующая на груз сила равна $-kx$. Но это верно только в статическом случае, если пружина растягивается достаточно медленно. Прн ускоренном движении груза (а следовательно, и пружины) пружина в разных своих частях растянута по-разному, и ее растяжение уже не пропорционально силе. При этом пружина уже не ведет себя квазистатически: она сама может колебаться как система с распределенными параметрами. Но если масса пружины мала по сравнению с массой прикрепленного к ней груза, то можно не считаться с этими колебаниямн, так как они «быстрые» по сравнению с колебаниями груза на пружине и очень быстро затухают.

В самом деле, частота колебаний груза, как видно из формулы (1.4), пропорциональна квадратному корню из отношения жесткости пружины к массе груза. При оценке частоты собственных колебаний пружины можно считать, что ее зависимость от жесткости пружины и массы имеет такой же вид. Поэтому при малой массе пружины частота колебаний велика по сравнению с частотой колебаний груза. Еслн для простоты предположить, что число колебаний за время их жизни одинаково по порядку величины как для колебаний груза, так и для колебаний самой пружины, то затухание высокочастотных колебаний пружины происходит за значительно меньшее время, чем затухание колебаний груза. Поэтому такие колебания пружины могли бы сыграть роль только в первый момент, когда они еще не затухли. Если в начальный момент пружина деформи-

рована однородно, то эти колебания вообще не возникают (разумеется, при условии, что масса пружины много меньше массы груза).

Если же в начальный момент пружина деформирована неоднородно, то такие быстрые колебания пружины как распределенной системы обязательно возникнут, но быстро затухнут, так что за время существования этих колебаний груз еще не успеет заметно сдвинуться с места. Что же может произойти в системе из-за этих колебаний?

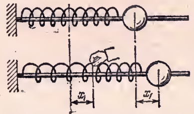


Рис. 3.1. В начальный момент растянута только левая половина пружины.

Проделаем такой опыт. Захватим пружину, изображенную на рис. 2.4, за середину и растянем ее левую половину на некоторую величину x_1 (рис. 3.1). Вторая половина пружины остается в недеформированном состоянии, так что груз в начальный момент смещен из положения равновесия вправо на величину x_1 и покоится. Затем пружину отпустим. К каким особенностям приведет то обстоятельство, что в начальный момент пружина деформирована неоднородно?

Если бы при смещении груза на x_1 пружина была деформирована однородно, то движение груза в отсутствие трения представляло бы собой гармоническое колебание около положения равновесия с частотой $\omega=\sqrt{k/m}$ и амплитудой x_1:

$$x(t)=x_1 \cos \omega t. \tag{3.1}$$

Начальная фаза колебаний в формуле (3.1) равна нулю, поскольку при $t=0$ груз смещен из положения равновесия на расстояние x_1, равное амплитуде колебаний. Однако в нашем случае пружина в начальный момент деформирована неоднородно — разные части пружины деформированы по-разному.

При однородной начальной деформации пружины запас механической энергии системы равен $kx_1^2/2$. При начальных условиях нашей задачи, когда растянута на x_1 половина пружины, запас энергии равен $2kx_1^2/2$, ибо, как нетрудно сообразить, жесткость «половины» пружины равна $2k$.

После затухания быстрых колебаний натяжение в пружине перераспределяется, а смещение груза остается равным x_1, так как груз за это время не успевает заметно сдвинуться. Деформация пружины становится однородной, а энергия системы становится равной $kx_1^2/2$. Таким образом, роль быстрых колебаний пружины свелась к тому, что запас энергии системы уменьшился до того значения, которое соответствует однородной начальной деформации пружины. Ясно, что дальнейшие процессы в системе не отличаются от случая однородной начальной деформации. Зависимость смещения груза от времени $x(t)$ выражается той же самой формулой (3.1). Напомним еще раз, что приведенное рассуждение справедливо при условии, что время затухания быстрых колебаний пружины много меньше периода колебаний груза на пружине.

В рассмотренном примере в результате быстрых колебаний превратилась во внутреннюю энергию (в тепло) половина первоначального запаса механической энергии. Ясно, что, подвергая начальной деформации не половину, а произвольную часть пружины, можно превратить во внутреннюю энергию любую долю первоначального запаса механической энергии. Но во всех случаях энергия колебаний груза на пружине соответствует запасу энергии при той же по величине однородной начальной деформации пружины.

Теперь представьте себе, что, не разобравшись в особенностях начальных условий, мы прямо применили бы закон сохранения механической энергии! Закон сохранения энергии универсален — мы много раз могли убедиться в этом. Но для его правильного применения нужна исчерпывающая «бухгалтерия»: необходимо тщательно разобраться, какие превращения энергии возможны в рассматриваемом явлении.

Таким образом, использованная при рассмотрении колебаний груза на пружине модель правильно описывает систему лишь в отсутствие колебаний пружины как распределенной системы. Несмотря на то, что эти колебания быстро прекращаются и не влияют на дальнейшее движение груза, они могут сильно отразиться на энергетических превращениях в системе.

А какие идеализации были сделаны при выводе уравнения (1.9) для колебаний в контуре? Чтобы разобраться

в этом, рассмотрим следующий пример, аналогичный разобранному выше случаю механических колебаний при неоднородной начальной деформации пружины.

В схеме, изображенной на рис. 3.2, левый конденсатор емкости C_1 имеет заряд q_0, а правый — емкости C_2 — не заряжен. Найдем колебания в контуре после замыкания ключа.

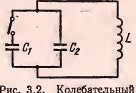


Рис. 3.2. Колебательный контур, эквивалентный механической системе на рис. 3.1.

Используя аналогию между механическими и электромагнитными колебаниями, мы можем представить себе картину происходящих в рассматриваемой системе процессов.

При замыкании ключа возникают быстрые затухающие колебания в контуре, состоящем из конденсаторов и соединяющих их проводов. Период таких колебаний очень мал, так как мала индуктивность соединительных проводов. В результате этих колебаний заряд на пластинах конденсаторов перераспределяется, после чего два конденсатора можно рассматривать как один. Но в первый момент, как и в рассмотренной механической системе, этого делать нельзя, ибо вместе с перераспределением зарядов происходит и перераспределение энергии, часть которой переходит в тепло.

После затухания быстрых колебаний в системе происходят колебания, как в контуре с одним конденсатором емкости C_1+C_2, заряд которого в начальный момент равен q_0. Как в механической системе необходимым условием справедливости приведенного рассуждения является малость массы пружины по сравнению с массой груза, так здесь — малость индуктивности соединительных проводов по сравнению с индуктивностью катушки.

ВЫНУЖДЕННЫЕ КОЛЕБАНИЯ

§ 4. Вынужденные колебания гармонического осциллятора. Резонанс

До сих пор мы рассматривали собственные колебания, т. е. колебания, происходящие в отсутствие внешних воздействий. Внешнее воздействие было нужно лишь для

того, чтобы вывести систему из состояния равновесия, после чего она предоставлялась самой себе. Уравнение собственных колебаний (1.5) или (1.9) вообще не содержит следов внешнего воздействия на систему: оно отражается лишь в начальных условиях.

Но очень часто приходится сталкиваться с колебаниями, которые происходят при постоянно присутствующем внешнем воздействии. Особенно важен и в то же время достаточно прост для изучения случай, когда внешняя сила имеет периодический характер. Общей чертой вынужденных колебаний, происходящих под действием периодической внешней силы, является то, что спустя некоторое время после начала действия внешней силы система полностью «забывает» свое начальное состояние, колебания приобретают стационарный характер и не зависят от начальных условий. Начальные условия проявляются только в период установления колебаний, который обычно называют переходным процессом.

Рассмотрим вначале наиболее простой случай установившихся вынужденных колебаний осциллятора под действием внешней силы, меняющейся по синусоидальному закону:

$$F(t)=F_0 \cos \omega t. \tag{4.1}$$

Такое внешнее воздействие на систему можно осуществить различными способами. Например, можно взять маятник в виде заряженного шарика на длинном стержне из диэлектрика и поместить его между вертикальными пластинами плоского конденсатора, на которые подается переменное синусоидальное напряжение (рис. 4.1). Вынужденные колебания маятника можно получить и чисто механическим путем. Для этого вместо конденсатора можно взять длинную пружину с малой жесткостью и прикрепить ее к стержню маятника недалеко от точки подвеса, как показано на рис. 4.2. Другой конец горизонтально расположенной пружины следует заставить двигаться по закону $B \cos \omega t$ с помощью кривошипно-шатунного механизма, приводимого в движение электромотором. Действующая на маятник со стороны пружины вынуждающая сила будет практически синусоидальна, если размах движения левого конца пружины B будет много больше амплитуды колебаний стержня маятника в точке закрепления пружины C.

Применяя к рассматриваемой системе второй закон Ньютона, можно убедиться, что уравнение движения маятника при не слишком больших амплитудах колебаний будет иметь вид

$$m\ddot{x} = -kx - \beta\dot{x} + F_0 \cos \omega t. \qquad (4.2)$$

Первое слагаемое в правой части представляет собой квазиупругую возвращающую силу, обусловленную действием

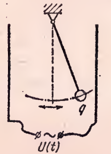


Рис. 4.1. Вынужденные колебания, возбуждаемые переменным электрическим полем.

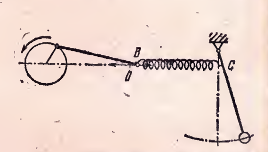


Рис. 4.2. Другой способ возбуждения вынужденных колебаний.

на маятник поля тяжести. Второе слагаемое есть сила трения, пропорциональная скорости, например сила сопротивления воздуха или сила трения в оси. Амплитуда вынуждающей силы F_0 равна произведению заряда маятника q на амплитуду переменной напряженности электрического поля E_0 в конденсаторе в первом случае и пропорциональна произведению максимального смещения левого конца пружины B на ее жесткость во втором случае.

Разделим обе части уравнения (4.2) на массу m и введем обозначения:

$$\omega_0^2 = \frac{k}{m}, \quad 2\gamma = \frac{\beta}{m}, \quad f_0 = \frac{F_0}{m}. \qquad (4.3)$$

Теперь уравнение (4.2) принимает вид

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = f_0 \cos \omega t. \qquad (4.4)$$

В отсутствие вынуждающей силы правая часть уравнения (4.4) обращается в нуль, и оно, как и следовало ожидать, сводится к уравнению собственных затухающих колебаний.

Вынужденные колебания в любой системе, способной в отсутствие трения совершать гармонические собственные колебания, будут при наличии синусоидальной вынуждающей силы и сопротивления, пропорционального скорости, описываться таким же уравнением (4.4). Величина x в любой системе характеризует отклонение от равновесия, а постоянные коэффициенты γ, ω_0 и f_0 определяются параметрами описываемой системы. Например, уравнение (4.4) описывает вынужденные колебания в контуре из последовательно соединенных емкости C, индуктивности L и сопротивления R, к которому приложено синусоидальное внешнее напряжение $U = U_0 \cos \omega t$ (рис. 4.3). Действительно, $U(t)$ в каждый момент времени равно сумме напряжений на отдельных элементах цепи:

Рис. 4.3. Последовательный RLC-контур.

$$U(t) = U_L + U_R + U_C. \qquad (4.5)$$

Подставляя в (4.5) $U_C = q/C$, $U_R = \dot{q}R$ и $U_L = L\ddot{q}$, получим уравнение

$$\ddot{q} + 2\gamma\dot{q} + \omega_0^2 q = f_0 \cos \omega t, \qquad (4.6)$$

где использованы обозначения:

$$\omega_0^2 = \frac{1}{LC}, \quad 2\gamma = \frac{R}{L}, \quad f_0 = \frac{U_0}{L}. \qquad (4.7)$$

Как найти решение уравнений (4.4) или (4.6), описывающее установившееся вынужденное колебание? При изучении переменного тока мы видели, что в последовательном контуре под действием синусоидального внешнего напряжения устанавливаются также синусоидальные колебания тока, происходящие с той же частотой, но с некоторым сдвигом по фазе относительно приложенного напряжения. Но эти колебания описываются уравнением (4.6). Значит, установившиеся вынужденные колебания в любой системе, описываемой таким уравнением, будут происходить по синусоидальному закону с частотой внешнего воздействия. Поэтому решение уравнений (4.6) или (4.4), которое

описывает не переходный процесс, а именно установившиеся колебания, следует искать в виде

$$x(t)=b\cos(\omega t-\theta), \qquad (4.8)$$

где частота ω совпадает с частотой в правой части уравнения (4.4), а постоянные амплитуда b и сдвиг фазы θ нужно выбрать так, чтобы функция (4.8) являлась решением уравнения (4.4). Амплитуду b и сдвиг фазы θ удобно определить с помощью векторных диаграмм, подобно тому как это делалось при изучении переменного тока.

Сопоставим каждому члену уравнения (4.4) вращающийся с угловой скоростью вектор, величина которого равна амплитудному значению этого члена. Мгновенное значение каждого члена получается проектированием соответствующего вектора на некоторое заранее выбранное направление. Поскольку проекция суммы нескольких векторов равна сумме проекций этих векторов, то уравнение (4.4) означает,

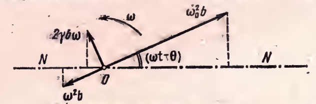


Рис. 4.4. Векторная диаграмма вынужденных колебаний, описываемых уравнением (4.4).

что сумма векторов, сопоставляемых членам, стоящим в левой части, равна вектору, сопоставляемому величине $f_0\cos\omega t$, стоящей в правой части. Чтобы построить эти векторы, выпишем мгновенные значения всех членов левой части уравнения (4.4), учитывая, что $x(t)$ дается формулой (4.8):

$$\begin{aligned}
\omega_0^2 x &= \omega_0^2 b\cos(\omega t-\theta),\\
2\gamma\dot{x} &= -2\gamma\omega b\sin(\omega t-\theta) = 2\gamma\omega b\cos\left(\omega t-\theta+\frac{\pi}{2}\right),\\
\ddot{x} &= -\omega^2 b\cos(\omega t-\theta) = \omega^2 b\cos(\omega t-\theta+\pi).
\end{aligned} \qquad (4.9)$$

Из формул (4.9) видно, что вектор длиной $2\gamma\omega b$, сопоставляемый величине $2\gamma\dot{x}$, опережает на угол $\pi/2$ вектор $\omega_0^2 b$, сопоставляемый величине $\omega_0^2 x$. Вектор $\omega^2 b$, сопоставляемый

члену $\ddot{x}$, опережает на π вектор $\omega_0^2 b$, т. е. эти векторы направлены в противоположные стороны. Взаимное расположение этих векторов для произвольного момента времени показано на рис. 4.4. Вся система векторов вращается как целое с угловой скоростью ω против часовой стрелки вокруг точки O. Мгновенные значения всех величин получаются проектированием соответствующих векторов на заранее

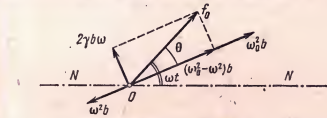


Рис. 4.5. Вектор f_0 сопоставляется правой части уравнения (4.4).

выбранное направление NN. Вектор, сопоставляемый правой части уравнения (4.4), равен сумме векторов, изображенных на рис. 4.4. Это сложение показано на рис. 4.5. Применяя теорему Пифагора, получим

$$f_0^2 = (\omega_0^2 - \omega^2)^2 b^2 + 4\gamma^2\omega^2 b^2,$$

откуда находим амплитуду установившихся вынужденных колебаний b:

$$b = \frac{f_0}{\sqrt{(\omega_0^2-\omega^2)^2 + 4\gamma^2\omega^2}}. \qquad (4.10)$$

Сдвиг фазы θ между вынуждающей силой $f(t)$ и смещением $x(t)$, как видно из векторной диаграммы на рис. 4.5, равен

$$\operatorname{tg}\theta = \frac{2\gamma\omega}{\omega_0^2-\omega^2}. \qquad (4.11)$$

Итак, установившиеся вынужденные колебания происходят по гармоническому закону (4.8), где b и θ определяются формулами (4.10) и (4.11).

Амплитуда установившихся вынужденных колебаний пропорциональна амплитуде вынуждающей силы f_0. Исследуем зависимость амплитуды колебаний от частоты ω вынуждающей силы. При малом затухании γ эта зависи-

мость носит очень резкий характер. Если $\gamma=0$, то при стремлении ω к частоте свободных колебаний ω_0 амплитуда вынужденных колебаний b стремится к бесконечности, как видно из формулы (4.10). При наличии затухания амплитуда колебаний в резонансе уже не обращается в бесконечность, хотя и значительно превышает амплитуду колебаний под действием внешней силы той же величины,

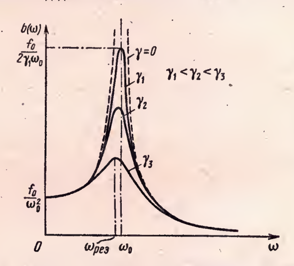


Рис. 4.6. Зависимость амплитуды установившихся вынужденных колебаний от частоты внешней силы.

но имеющей частоту, далекую от резонансной. Резонансные кривые при разных значениях постоянной затухания γ приведены на рис. 4.6. Для нахождения частоты резонанса $\omega_{рез}$ нужно найти, при каком ω подкоренное выражение в формуле (4.10) имеет минимум. Приравнивая производную этого выражения по ω нулю (или дополняя его до полного квадрата), убеждаемся, что максимум амплитуды вынужденных колебаний имеет место при

$$\omega_{рез}=\sqrt{\omega_0^2-2\gamma^2}. \qquad (4.12)$$

Резонансная частота оказывается меньше частоты свободных колебаний системы. При малых γ резонансная частота практически совпадает с ω_0.

При стремлении частоты вынуждающей силы к бесконечности, т. е. при $\omega \gg \omega_0$, амплитуда b, как видно из (4.10), стремится к нулю. При $\omega=0$, т. е. при действии постоянной внешней силы, величина b равна f_0/ω_0^2. Если подставить сюда $f_0=F_0/m$ и $\omega_0^2=k/m$, получим $b_{ст}=F_0/k$. Это есть статическое смещение осциллятора из положения равновесия под действием постоянной силы F_0. Амплитуду вынужденных колебаний в резонансе $b_{рез}$ находим, подставляя частоту $\omega_{рез}$ из (4.12) в выражение (4.10):

$$b_{рез}=\frac{f_0}{2\gamma\sqrt{\omega_0^2-\gamma^2}} \approx \frac{f_0}{2\gamma\omega_0}. \qquad (4.13)$$

Амплитуда колебаний в резонансе тем больше, чем меньше постоянная затухания γ. При изучении вынужденных колебаний вблизи резонанса трением пренебрегать нельзя, как бы мало оно ни было: только при учете затухания амплитуда в резонансе $b_{рез}$ получается конечной. Интересно сравнить значение $b_{рез}$ со статическим смещением $b_{ст}$ под действием силы F_0. Составляя отношение $b_{рез}/b_{ст}$, получим при малом затухании

$$\frac{b_{рез}}{b_{ст}}=\frac{\omega_0}{2\gamma}. \qquad (4.14)$$

Подставляя сюда $\omega_0=2\pi/T$ и учитывая, что $1/\gamma=\tau$ есть время жизни собственных затухающих колебаний для той же системы в отсутствие внешних сил, находим

$$\frac{b_{рез}}{b_{ст}}=\pi\frac{\tau}{T}.$$

Но τ/T есть число колебаний, совершаемых затухающим осциллятором за время жизни колебаний τ. Таким образом, резонансные свойства системы характеризуются тем же параметром, что и собственные затухающие колебания.

Формула (4.11) дает возможность проанализировать изменение сдвига фазы θ между внешней силой и смещением $x(t)$ при вынужденных колебаниях. При $\omega \ll \omega_0$ значение $\operatorname{tg}\theta$ близко к нулю. Это означает, что при низких частотах смещение осциллятора происходит в фазе с внешней силой. При медленном вращении кривошипа на рис. 4.2 маятник движется в такт с правым концом шатуна. Если $\omega \gg \omega_0$,

то $\operatorname{tg}\theta$ стремится к нулю со стороны отрицательных значений, т. е. сдвиг фазы равен π, и смещение осциллятора происходит в противофазе с вынуждающей силой. В резонансе, как видно из формулы (4.11), смещение отстает по фазе от внешней силы на $\pi/2$. Вторая из формул (4.9) показывает, что при этом внешняя сила изменяется в фазе со скоростью $\dot{x}(t)$, т. е. все время действует в направлении

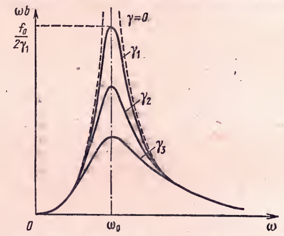


Рис. 4.7. Амплитуда скорости при установившихся вынужденных колебаниях.

движения. Что именно так и должно быть, ясно и из интуитивных соображений.

Из этой же формулы (4.9) видно, что амплитуда колебаний скорости при установившихся вынужденных колебаниях равна ωb. С помощью (4.10) получаем

$$\omega b=\frac{f_0\omega}{\sqrt{(\omega_0^2-\omega^2)^2+4\gamma^2\omega^2}}=\frac{f_0}{\sqrt{\left(\frac{\omega_0^2}{\omega}-\omega\right)^2+4\gamma^2}}. \quad (4.15)$$

Зависимость амплитуды скорости от частоты внешней силы показана на рис. 4.7. Резонансная кривая для скорости хотя и похожа на резонансную кривую для смещения, но отличается от нее в некоторых отношениях. Так, при $\omega=0$, т. е. при действии постоянной силы, осциллятор испытывает статическое смещение из положения равновесия, и скорость

его после того, как закончится переходный процесс, равна нулю. Из формулы (4.15) видно, что амплитуда скорости при $\omega=0$ обращается в нуль. Резонанс скорости имеет место при точном совпадении частоты внешней силы с частотой свободных колебаний ω_0.

Обратим внимание, что при изучении резонансных явлений в цепи переменного тока при действии синусоидального внешнего напряжения исследовались вынужденные колебания тока в цепи $I=\dot{q}$, который является аналогом скорости $\dot{x}$ (а не смещения x) в случае механических колебаний. Поэтому все сказанное в этом параграфе о скорости $\dot{x}(t)$ при вынужденных колебаниях справедливо и для тока $I(t)$ в последовательном контуре, а все сказанное о смещении $x(t)$ справедливо для заряда конденсатора $q(t)$.

§ 5. Энергетические превращения при вынужденных колебаниях. Установление колебаний

Установившиеся вынужденные колебания под действием синусоидальной силы внешне очень похожи на собственные незатухающие колебания: они происходят по синусоидальному закону с неизменной амплитудой. Но, несмотря на внешнее сходство, это принципиально разные колебания. При свободных колебаниях энергия колебаний, т. е. сумма кинетической и потенциальной энергий, постоянна, а средние значения кинетической и потенциальной энергий равны между собой. А как обстоит дело в случае синусоидальных вынужденных колебаний? Запишем выражение для энергии колебаний:

$$E=\frac{1}{2}m\dot{x}^2+\frac{1}{2}kx^2=$$
$$=\frac{1}{2}m\omega^2b^2\sin^2(\omega t-\theta)+\frac{1}{2}kb^2\cos^2(\omega t-\theta). \quad (5.1)$$

Входящая в это выражение частота ω определяется внешним воздействием и не зависит от характеризующих осциллятор величин k и m. Поэтому в отличие от случая свободных колебаний, где $\omega_0^2=k/m$ и множители перед $\sin^2(\omega_0 t+\alpha)$ и $\cos^2(\omega_0 t+\alpha)$ в формуле (1.16) оказались одинаковыми, в формуле (5.1) это не так. Таким образом, полная энергия при установившихся вынужденных коле-

баниях непостоянна. На рис. 5.1 показана зависимость от времени кинетической, потенциальной и полной энергии осциллятора при установившихся вынужденных колебаниях в случае $\omega \ll \omega_0$. Все время идет переход энергии от источника внешнего воздействия в рассматриваемую систему и

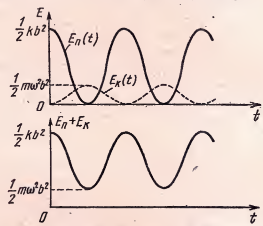


Рис. 5.1. Кинетическая, потенциальная и полная энергия осциллятора при установившихся вынужденных колебаниях ($\omega \ll \omega_0$).

обратно. Полная энергия постоянна только при $\omega = \omega_0$, т. е. при резонансе.

Средние значения кинетической и потенциальной энергий при вынужденных колебаниях могут сильно отличаться друг от друга. При низких частотах, когда $\omega < \omega_0$, среднее значение кинетической энергии меньше среднего значения потенциальной; при $\omega > \omega_0$ — наоборот. Действительно, при вынужденных колебаниях с очень низкими частотами почти вся энергия осциллятора — это энергия деформированной пружины, а кинетическая энергия ничтожно мала. При высоких частотах, напротив, скорость может достигать огромных значений даже при ничтожных смещениях, когда потенциальная энергия пренебрежимо мала.

Рассмотрим подробнее энергетические превращения при установившихся вынужденных колебаниях. Если частота внешней силы много меньше частоты собственных колебаний системы, то, как уже отмечалось, почти вся энергия

колебаний представляет собой потенциальную энергию. Поэтому, когда осциллятор удаляется от положения равновесия, энергия системы возрастает, т. е. внешняя сила совершает положительную работу. На протяжении этой четверти периода энергия поступает в систему от внешнего источника. На протяжении следующей четверти периода, когда осциллятор возвращается в положение равновесия и потенциальная энергия убывает, система отдает энергию внешнему источнику. Затем все повторяется.

Если частота внешней силы много больше частоты собственных колебаний, то, как мы видели, энергия осциллятора — это в основном кинетическая энергия. Поэтому система получает энергию от внешнего источника в те четверти периода, когда осциллятор движется к положению равновесия и его скорость возрастает. При удалении от положения равновесия система отдает энергию внешнему источнику.

Ясно, что при установившихся колебаниях получаемая системой от внешнего источника за период энергия превосходит отдаваемую, так как в системе действует сила трения, работа которой определяет диссипацию механической энергии — переход части энергии колебаний в тепло.

При резонансе, когда частота внешней силы совпадает с частотой свободных колебаний, полная энергия системы постоянна, как и в случае свободных колебаний. Дважды за период кинетическая и потенциальная энергии целиком переходят друг в друга. Другими словами, при резонансе система совершает «почти собственные» колебания. Роль внешней силы сводится только к компенсации действующей в системе силы трения.

Запишем выражение для развиваемой внешней силой мощности $P(t)$ при установившихся колебаниях:

$$P(t)=F(t)\,v_x(t)=F(t)\,\dot{x}(t).$$

Подставляя сюда выражения для внешней силы и скорости осциллятора, получим

$$P(t)=-F_0\cos\omega t\cdot b\omega\sin(\omega t-\theta).$$

Используя формулу тригонометрии для произведения синуса и косинуса, приводим выражение для мощности к

следующему виду:

$$P(t) = -\frac{1}{2} F_0 b\omega \left[\sin(2\omega t - \theta) - \sin\theta\right]. \qquad (5.2)$$

Поскольку среднее значение $\sin(2\omega t - \theta)$ за период равно нулю, то среднее за период значение мощности внешней силы $\langle P \rangle$ равно

$$\langle P \rangle = \frac{1}{2} F_0 b\omega \sin\theta. \qquad (5.3)$$

Обратим внимание на то, что эта формула совершенно аналогична выражению для средней мощности в цепи переменного тока. Отличие лишь в том, что вместо амплитуд приложенного напряжения U_0 и тока I_0 стоят амплитуды внешней силы F_0 и скорости $b\omega$. Вместо привычного $\cos\varphi$ стоит $\sin\theta$ потому, что здесь мощность выражена через сдвиг фазы θ между силой и смещением, а не между силой и скоростью, обозначавшийся через φ.

Скорость диссипации механической энергии в системе определяется мощностью, развиваемой силой трения:

$$P_{\text{тр}}(t) = F_{\text{тр}} v_x(t) = -\beta v_x^2(t) = -\beta \dot{x}^2(t). \qquad (5.4)$$

Подставляя сюда значение скорости $\dot{x}(t)$, находим

$$P_{\text{тр}}(t) = -\beta b^2\omega^2 \sin^2(\omega t - \theta) = -\frac{1}{2}\beta\omega^2 b^2\left[1 - \cos 2(\omega t - \theta)\right]. \qquad (5.5)$$

Усредняя по периоду колебаний, получим

$$\langle P_{\text{тр}} \rangle = -\frac{1}{2}\beta b^2\omega^2. \qquad (5.6)$$

При установившихся колебаниях средняя энергия осциллятора постоянна. Получаемая системой от внешнего источника за период энергия полностью превращается в тепло. Это значит, что равна нулю сумма средней мощности внешней силы и силы трения:

$$\langle P \rangle + \langle P_{\text{тр}} \rangle = 0. \qquad (5.7)$$

Линейная зависимость средней мощности внешней силы и квадратичная зависимость мощности силы трения от амплитуды колебаний позволяют объяснить устойчивость режима вынужденных колебаний. Изобразим эти зависи-

мости графически. На рис. 5.2 прямая линия характеризует получаемую системой энергию, а парабола — диссипируемую энергию, определяемую мощностью силы трения. Поскольку в установившемся режиме эти энергии равны, то точка пересечения прямой и параболы соответствует амплитуде установившихся колебаний. Представим себе, что в силу каких-то случайных причин амплитуда колебаний немного изменилась, например уменьшилась при неизменной фазе. Тогда, как видно из рис. 5.2, мощность внешней силы будет больше диссипируемой мощности. Это приводит к росту энергии системы и восстановлению прежнего значения амплитуды колебаний.

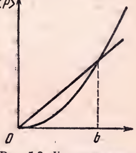


Рис. 5.2. К исследованию устойчивости режима вынужденных колебаний.

Аналогично можно убедиться в том, что амплитуда вынужденных колебаний устойчива и по отношению к случайным отклонениям в сторону возрастания.

До сих пор мы рассматривали установившийся режим вынужденных колебаний. А как происходит установление колебаний?

Пусть в начальный момент осциллятор покоится в положении равновесия, т. е. начальные условия имеют вид

$$x(0)=0, \quad \dot{x}(0)=0. \tag{5.8}$$

В этот момент на него начинает действовать внешняя синусоидальная сила на частоте ω, равной частоте ω_0 свободных колебаний осциллятора. Как мы знаем, движение осциллятора будет описываться уравнением (4.4):

$$\ddot{x}+2\gamma\dot{x}+\omega_0^2x=f_0\cos\omega_0 t. \tag{5.9}$$

Нам известно решение этого уравнения, описывающее установившиеся колебания, которые не зависят от начальных условий. Оно, согласно (4.8), имеет вид

$$x(t)=b\cos\left(\omega_0 t-\frac{\pi}{2}\right)=b\sin\omega_0 t. \tag{5.10}$$

В этом выражении сдвиг фазы θ между силой и смещением положен равным $\pi/2$, так как частота вынужденных коле-

баний в рассматриваемом примере равна частоте свободных колебаний ω_0. Однако это решение не удовлетворяет начальным условиям (5.8), так как, согласно (5.10), скорость $\dot{x}$ при $t=0$ не равна нулю. Как же найти решение уравнения (5.9), удовлетворяющее нашим начальным условиям? Такое решение обязательно должно переходить в (5.10) по мере установления колебаний, т. е. при $t\to\infty$. Поэтому попробуем искать решение в виде суммы выражения (5.10) и функции $Ae^{-\gamma t}\cos(\omega_0 t+\alpha)$, описывающей собственные затухающие колебания осциллятора, т. е. являющейся решением уравнения (5.9) с правой частью, равной нулю, в случае малого затухания $\gamma\ll\omega_0$. Такая сумма

$$x(t)=b\sin\omega_0 t+Ae^{-\gamma t}\cos(\omega_0 t+\alpha) \qquad (5.11)$$

действительно является решением уравнения (5.9), в чем можно убедиться непосредственной подстановкой. В самом деле, уравнение (5.9) содержит функцию $x(t)$ и ее производные только в первой степени, поэтому каждое слагаемое в выражении (5.11) можно подставлять в уравнение (5.9) по отдельности. Подстановка слагаемого $b\sin\omega_0 t$ в левую часть (5.9) дает $f_0\cos\omega_0 t$, а подстановка второго слагаемого дает нуль.

Благодаря множителю $e^{-\gamma t}$ второе слагаемое в (5.11) стремится к нулю при $t\to\infty$, и остается только член $b\sin\omega_0 t$, описывающий установившиеся вынужденные колебания. Но при малых значениях времени t второе слагаемое в (5.11) играет важную роль: наличие двух произвольных постоянных A и α позволяет удовлетворить любым начальным условиям. Полагая в (5.11) $t=0$ и учитывая первое из начальных условий (5.8), получим

$$0=A\cos\alpha,$$

откуда $\alpha=\pi/2$ и $\cos(\omega_0 t+\alpha)$ в (5.11) равен $-\sin\omega_0 t$. При нахождении скорости $\dot{x}$ из (5.11) учтем, что при малом затухании, когда $\gamma\ll\omega_0$, сомножитель $e^{-\gamma t}$ почти не изменяется на протяжении периода колебаний. Поэтому при дифференцировании $x(t)$ его можно считать постоянным:

$$\dot{x}(t)=b\,\omega_0\cos\omega_0 t-Ae^{-\gamma t}\omega_0\cos\omega_0 t. \qquad (5.12)$$

Полагая здесь $t=0$ и учитывая второе начальное условие (5.8), получаем

$$0=b\omega_0-A\,\omega_0,$$

откуда $A=b$. Теперь выражение (5.11) принимает вид

$$x(t)=b\sin\omega_0 t-be^{-\gamma t}\sin\omega_0 t=b(1-e^{-\gamma t})\sin\omega_0 t. \quad (5.13)$$

Первое слагаемое в (5.13) $b\sin\omega_0 t$ представляет собой гармоническое колебание постоянной амплитуды и соответствует установившимся вынужденным колебаниям. Второе

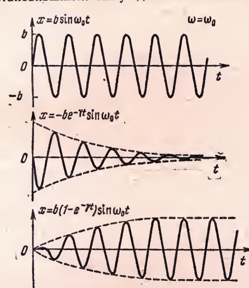


Рис. 5.3. Процесс установления вынужденных колебаний при резонансе.

слагаемое соответствует собственным затухающим колебаниям. Поэтому процесс установления колебаний можно представить себе таким образом: в начале процесса в системе одновременно присутствуют и вынужденные, и собственные колебания, причем амплитуда и фаза последних таковы, чтобы результирующее колебание удовлетворяло начальным условиям. Графики этих колебаний показаны на рис. 5.3

При малом затухании результирующее колебание $x(t)$ в (5.13) можно рассматривать как синусоидальное колебание с частотой ω_0, амплитуда которого медленно нарастает со временем (рис. 5.3). Характерное время

установления амплитуды колебаний $\tau=1/\gamma$ совпадает со временем жизни собственных затухающих колебаний в той же системе.

Подведем некоторые итоги. При очень малом затухании амплитуда в резонансе будет очень большой, но ее установление длится очень долго. Чем более резко выражен

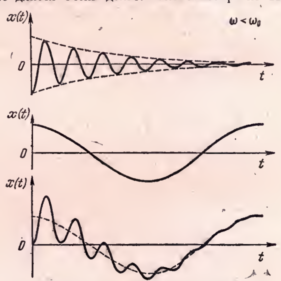


Рис. 5.4. Установление вынужденных колебаний при $\omega<\omega_0$.

резонанс, тем медленнее происходит установление. Это легко понять и с помощью энергетических соображений: чем острее резонанс, тем больше запасаемая системой энергия и, следовательно, тем больше времени требуется для того, чтобы сообщить системе эту энергию.

Если частота вынуждающей силы ω не совпадает с частотой свободных колебаний ω_0, то процесс установления колебаний также можно представить как наложение вынужденных колебаний с частотой ω и затухающих собственных колебаний с частотой ω_0. Картина установления колебаний при $\omega<\omega_0$ показана на рис. 5.4.

Вынужденные колебания осциллятора возможны при любом периодическом внешнем воздействии, а не только синусоидальном. При этом установившееся колебание,

вообще говоря, не будет синусоидальным, но оно будет представлять собой периодическое движение с периодом, равным периоду внешнего воздействия. Внешнее воздействие $F(t)$ может представлять собой, например, последовательность периодически повторяющихся толчков

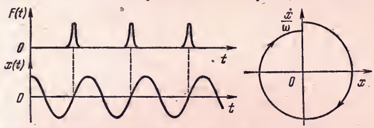


Рис. 5.5. Внешнее воздействие на осциллятор имеет вид коротких толчков, частота следования которых совпадает с собственной частотой осциллятора.

Рис. 5.6. Фазовая диаграмма колебаний, происходящих под действием очень коротких толчков.

(рис. 5.5). Если период внешних толчков совпадает с периодом собственных колебаний, то в системе наступает резонанс. Колебания при этом будут почти синусоидальными. Сообщаемая системе при каждом толчке энергия при резонансе мала по сравнению с запасом энергии системы и равна диссипируемой за период энергии.

На рис. 5.6 показана фазовая диаграмма вынужденных колебаний осциллятора, происходящих под действием коротких толчков. При каждом толчке осциллятор изменяет скорость на одну и ту же величину Δv. Период чередования толчков равен периоду собственных колебаний осциллятора, т. е. имеет место резонанс. Движение осциллятора установится таким образом, что толчки будут приходиться на те моменты времени, когда осциллятор проходит положение равновесия.

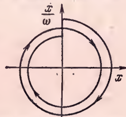


Рис. 5.7. Период следования толчков вдвое превосходит период осциллятора.

Резонанс будет иметь место и в том случае, когда период чередования толчков будет кратен периоду собственных

колебаний. Такое невозможно при синусоидальном внешнем воздействии. На рис. 5.7 показана фазовая диаграмма для случая, когда период толчков вдвое превышает период осциллятора.

§ 6. Автоколебания

Наиболее интересными, хотя и очень сложными для исследования являются системы, в которых колебания возникают не за счет начального толчка и не за счет периодического внешнего воздействия, а в результате имеющейся у каждой из таких систем способности самой регулировать поступление энергии от постоянного источника. Такие системы носят название автоколебательных. Наиболее известный пример автоколебательной системы — обычный часовой механизм.

Характерные элементы автоколебательной системы, или, как говорят, генератора незатухающих колебаний,— это резонатор, источник энергии и обратная связь между резонатором и источником энергии. Резонатор представляет собой систему, в которой могут происходить собственные затухающие колебания. Примерами резонатора могут служить маятник настенных часов или балансир наручных, колебательный контур в ламповом генераторе, струна в смычковом музыкальном инструменте. Обратная связь представляет собой устройство, с помощью которого генератор сам регулирует поступление энергии от источника.

Обратная связь в приведенных примерах осуществляется анкерным механизмом в часах, индуктивно связанной с колебательным контуром катушкой обратной связи в ламповом генераторе, смычком в руках музыканта.

Наличие трения в резонаторе приводит к диссипации энергии колебаний. Однако обратная связь обеспечивает необходимое восполнение энергии, так что амплитуда колебаний остается неизменной. Как и при вынужденных колебаниях под действием периодической внешней силы, при автоколебаниях, независимо от начального состояния, в конце концов устанавливается стационарный режим колебаний с определенной частотой и амплитудой. Но в отличие от установившихся вынужденных колебаний, где частота и амплитуда определяются внешним воздействием, в случае

автоколебаний как частота, так и амплитуда определяются только свойствами самой системы.

Анализ уравнений, описывающих даже самые простые реальные автоколебательные системы, представляет собой сложную задачу. Поэтому мы рассмотрим упрощенную модель автоколебательной системы, допускающую сравнительно простое исследование.

Ранее был рассмотрен пример колебаний, затухающих под действием силы сухого трения, величина которой не зависит от скорости. Пока направление скорости оставалось неизменным, движение при наличии постоянной силы трения происходило так, как при незатухающем гармоническом колебании около смещенного положения равновесия (рис. 2.5). При перемене направления скорости и, следовательно, направления силы трения происходило изменение положения равновесия. Затухание колебаний, как видно из фазовой диаграммы на рис. 2.6, проявлялось в переходах от одного положения равновесия к другому, хотя диссипация энергии и выделение тепла при наличии сухого трения происходили непрерывно.

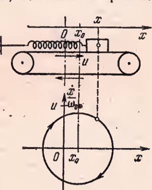


Рис. 6.1. Колебания груза на движущейся ленте транспортера при наличии сухого трения.

А что будет, если каким-либо способом добиться того, чтобы перескоков от одного положения равновесия к другому не происходило? Этого легко добиться, например, положив прикрепленный к пружине брусок на ленту транспортера, которая движется со скоростью u, большей максимального значения скорости колебаний груза на пружине (рис. 6.1). Поскольку в этом случае на брусок все время действует постоянная сила трения μmg, направленная направо, то уравнение движения бруска имеет вид

$$m\ddot{x} = -kx + \mu mg. \qquad (6.1)$$

Это уравнение описывает гармонические колебания с частотой $\omega_0 = \sqrt{k/m}$ около нового положения равновесия

x_0, в котором пружина находится в растянутом на величину $x_0=\mu mg/k$ состоянии. Тогда уравнение (6.1) можно записать в виде

$$\ddot{x} = -\omega_0^2 (x - x_0). \tag{6.2}$$

Это значит, что координата бруска x зависит от времени по следующему закону:

$$x(t) = x_0 + A \cos \omega_0 t. \tag{6.3}$$

В выражении (6.3) начало отсчета времени выбрано в момент наибольшего смещения бруска направо. Скорость движения бруска $\dot{x}$ в этом случае равна

$$\dot{x}(t) = -A\omega_0 \sin \omega_0 t. \tag{6.4}$$

Фазовая траектория этого движения изображена на нижней части рис. 6.1. В координатах x, $\dot{x}/\omega_0$ она представляет собой окружность с центром в точке x_0. Изображающая состояние бруска точка движется вдоль фазовой траектории по часовой стрелке. Радиус окружности равен амплитуде колебаний A, и он, как было указано, меньше величины u/ω_0, где u — скорость ленты транспортера.

Рассмотрим описанные колебания груза с точки зрения энергетических превращений в системе. Прежде всего вспомним, что в отсутствие трения фазовая траектория свободных колебаний представляет собой окружность с центром в начале координат (рис. 1.7). Так как энергия в такой системе постоянна, то всем точкам этой окружности соответствует одна и та же энергия. Энергия осциллятора пропорциональна квадрату радиуса окружности. Чем дальше лежит изображающая состояние точка от начала координат, тем больше энергия системы в этом состоянии.

Теперь из фазовой диаграммы на рис. 6.1 видно, что на нижней половине окружности, соответствующей движению бруска справа налево, энергия системы убывает, а на верхней полуокружности, т. е. при движении слева направо, энергия возрастает. Эти изменения энергии обусловлены исключительно действием силы трения. Пока брусок движется справа налево, сила трения направлена против скорости и тормозит движение бруска, уменьшая энергию системы. Но при движении слева направо сила

трения направлена вдоль скорости и, подталкивая брусок, увеличивает энергию системы. Это и приводит к возможности существования незатухающих колебаний в системе: убыль энергии за одну половину периода восполняется за другую половину. И все это происходит в результате действия постоянной силы трения.

Приведенные результаты можно проиллюстрировать и с помощью закона сохранения энергии. Действительно, полная энергия бруска на пружине при рассматриваемых колебаниях равна

$$E = E_к + E_п = \frac{1}{2} m\dot{x}^2 + \frac{1}{2} kx^2 =$$
$$= \frac{1}{2} m\omega_0^2 A^2 \sin^2 \omega_0 t + \frac{1}{2} k (x_0 + A \cos \omega_0 t)^2.$$

Раскрывая скобки в последнем слагаемом и перегруппировывая члены, найдем

$$E = \frac{1}{2} k (x_0^2 + A^2) + kx_0 A \cos \omega_0 t. \qquad (6.5)$$

Выражение для энергии (6.5) содержит осциллирующее слагаемое, среднее значение которого за период равно нулю. Из (6.5) легко найти скорость изменения энергии dE/dt:

$$\frac{dE}{dt} = -kx_0 A\omega_0 \sin \omega_0 t. \qquad (6.6)$$

В соответствии с законом сохранения энергии эта величина должна быть равна мощности действующей на брусок силы трения. Непосредственно вычисляя мощность силы трения:

$$P_{тр} = F_{тр} \dot{x}(t) = -kx_0\omega_0 A \sin \omega_0 t, \qquad (6.7)$$

убеждаемся, что она совпадает с правой частью выражения (6.6). Таким образом, мощность силы трения на протяжении периода колебаний принимает и отрицательные, и положительные значения, а ее работа за период равна нулю: по истечении целого периода энергия системы принимает прежнее значение.

Если бы лента транспортера была неподвижной, то, как видно из формулы (6.7), мощность силы трения была бы

все время отрицательной, так как сила трения меняла бы свое направление одновременно с изменением направления скорости. Это приводило бы к быстрому прекращению колебаний.

Остановимся на особенностях выделения тепла в результате трения в рассматриваемом примере. Сила трения, действующая со стороны бруска на ленту транспортера, постоянна и равна μmg, или kx_0. Поэтому приводящий ленту в движение со скоростью u мотор развивает постоянную мощность P, равную kx_0u. Если бы брусок не совершал колебаний, то выделяющееся в единицу времени тепло было бы равно этой мощности. При колебаниях же бруска только среднее значение скорости выделения тепла равно kx_0u. В те половины периодов колебаний, когда действующая на брусок сила трения увеличивает энергию осциллятора, скорость выделения тепла меньше этого среднего значения kx_0u. В те половины периодов, когда сила трения уменьшает энергию осциллятора, скорость выделения тепла превосходит мощность, развиваемую мотором. Все это отчетливо видно из выражения для скорости выделения тепла dQ/dt, которая определяется относительной скоростью ленты и бруска $u - \dot{x}$:

$$\frac{dQ}{dt} = F_{\text{тр}}(u - \dot{x}) = kx_0(u + \omega_0 A \sin \omega_0 t). \qquad (6.8)$$

Подчеркнем, что dQ/dt всегда положительна, поскольку скорость ленты u больше амплитудного значения скорости $\omega_0 A$. Так, разумеется, и должно быть, так как при трении тепло может только выделяться.

Рассмотрим теперь, как происходит установление стационарного режима колебаний в изучаемой модели автоколебательной системы и чем определяется значение амплитуды. Начнем с наиболее простого случая: брусок в начальный момент покоится в положении, соответствующем недеформированной пружине. Точка, изображающая такое начальное состояние, находится в начале координат на фазовой плоскости. Пусть скорость ленты u настолько велика, что $u > x_0\omega_0$. Сила трения начинает ускорять брусок, и его движение происходит, согласно уравнению (6.2), так же, как и при гармоническом колебании около точки x_0. В течение первой четверти периода с момента начала движения скорость бруска растет, но все же вследствие

условия $u > x_0\omega_0$ не достигает значения, равного скорости ленты. Действительно, фазовая траектория этого движения представляет собой окружность с центром x_0, проходящую через начальное состояние, т. е. через начало координат (рис. 6.2). Радиус этой окружности, равный x_0, и представляет собой амплитуду автоколебаний A. Так как этот радиус, согласно условию $u > x_0\omega_0$, меньше u/ω_0, то скорость ленты больше амплитудного значения скорости бруска и сила трения все время направлена в одну сторону.

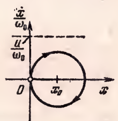


Рис. 6.2. Фазовая диаграмма в случае быстрого движения ленты (или малого коэффициента трения).

Пусть теперь скорость ленты такова, что $u < x_0\omega_0$. Теперь фазовая траектория на начальном участке движения бруска будет представлять собой окружность с центром в точке x_0 только до тех пор, пока скорость бруска не сравняется со скоростью ленты u. Этому моменту соответствует точка B фазовой диаграммы на рис. 6.3. В этот момент проскальзывание прекращается, трение скольжения заменяется трением покоя и величина силы трения скачком падает до значения kx_1, равного силе натяжения пружины. Брусок движется вместе с лентой с постоянной скоростью u, растягивая пружину до тех пор, пока сила натяжения пружины kx не станет равной максимальному значению силы трения покоя $\mu mg = kx_0$. С этого момента снова начинается проскальзывание, и фазовая траектория дальнейшего движения бруска представляет собой окружность с центром в точке x_0. Ее радиус равен u/ω_0. Это и есть значение амплитуды A установившихся автоколебаний.

Такая «сшитая» из различных кусков фазовая траектория всегда соответствует тому, что движение тела на разных участках описывается разными уравнениями. В самом деле, движение на начальном участке OB происходит в соответствии с уравнением колебаний (6.2). На участке BC, где сила трения покоя уравновешивает силу натяжения пружины, уравнение движения имеет вид $\ddot{x} = 0$, а его решение есть $\dot{x} = u$. В дальнейшем движение снова описывается уравнением колебаний (6.2). На нижней части

рис. 6.3 приведена зависимость смещения бруска от времени для рассмотренного случая.

Рассмотрим, как происходит установление колебаний при других начальных условиях. Пусть, например, находящемуся в точке $x=0$ бруску сообщается начальная скорость, бо́льшая скорости ленты u. Такому начальному состоянию соответствует точка B на рис. 6.4. В этом случае в начальный момент действующая на брусок сила трения скольжения направлена налево и уравнение движения имеет вид

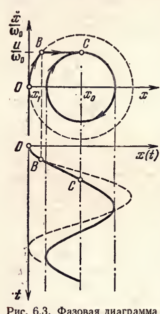


Рис. 6.3. Фазовая диаграмма и график смещения в случае $u<x_0\omega_0$.

$$m\ddot{x} = -kx - \mu mg. \quad (6.9)$$

Это уравнение описывает гармоническое колебание с частотой

$$\omega_0 = \sqrt{\frac{k}{m}}$$

около положения равновесия $-x_0$, где $x_0=\mu mg/k$. Поэтому начальный участок фазовой траектории представляет собой часть окружности с центром в точке $-x_0$. Как только скорость бруска уменьшится до значения, равного скорости ленты (точка C на рис. 6.4), сила трения скачком меняет направление на противоположное. Дальнейшее движение бруска представляет собой колебание около положения равновесия x_0. На фазовой диаграмме ему соответствует часть окружности с центром в точке x_0. В точке D, где скорость бруска снова станет равна скорости ленты u, может произойти следующее. Если точка D окажется правее точки $-x_0$, то в этот момент трение скольжения заменится трением покоя, и дальше все будет происходить так же, как на рис. 6.3. Этот случай показан на рис. 6.4. Если же точка D окажется левее точки $-x_0$, то сила трения скольжения опять скачком изменит направление, и фазовая траектория дальнейшего движения

будет представлять собой часть окружности с центром в точке $-x_0$. И так будет продолжаться до тех пор, пока фазовая траектория не попадет на прямую $\dot{x}=u$ в промежутке между точками $-x_0$ и x_0. После этого фазовая траектория, как и прежде, выйдет на ту же самую предельную окружность с центром в точке x_0 и радиусом $A=u/\omega_0$.

Таким образом, если начальное состояние системы изображается любой точкой на фазовой плоскости, лежащей

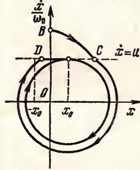


Рис. 6.4. Фазовая диаграмма в случае, когда начальная скорость бруска больше скорости ленты.

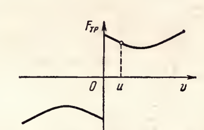


Рис. 6.5. Зависимость силы сухого трения от скорости.

вне этой предельной окружности, то фазовая траектория системы в конце концов выходит на нее. Амплитуда установившихся автоколебаний, т. е. радиус предельной окружности, как мы видим, при этом не зависит от начального состояния.

Если же начальное состояние изображается точкой, лежащей внутри предельной окружности, то фазовая траектория колебаний представляет собой окружность с центром в точке x_0, проходящую через начальное состояние. Такой случай показан на рис. 6.2.

Может показаться, что рассмотренная упрощенная модель автоколебательной системы на самом деле не будет работать, если ее осуществить на опыте. Дело в том, что в любой системе всегда присутствует трение, пропорциональное скорости, например сопротивление воздуха. В результате на фазовой плоскости вместо предельного цикла — окружности с центром в x_0 — будет получаться спираль,

закручивающаяся вокруг точки x_0, подобно тому как это происходило при собственных затухающих колебаниях. Так бы, разумеется, и было, если бы сила сухого трения действительно совершенно не зависела от величины скорости. Но реально небольшая зависимость все-таки есть. Она показана на рис. 6.5. Осуществить модель автоколебательной системы удается потому, что на графике этой зависимости имеется «падающий» участок, где трение убывает с увеличением скорости. Если выбрать скорость ленты u так, как показано на рис. 6.5, то можно добиться компенсации трения, зависящего от скорости, и осуществить предельный цикл в такой системе, т. е. получить незатухающие автоколебания.

При практическом выполнении опыта удобно использовать не брусок на движущейся ленте, а маятник в виде жесткого стержня с муфтой, надетой на вращающийся вал.

§ 7. Несинусоидальные колебания

Во всех упоминавшихся выше примерах автоколебательных систем обязательным элементом являлся резонатор. Другими словами, в отсутствие обратной связи в этих системах возможны собственные затухающие колебания.

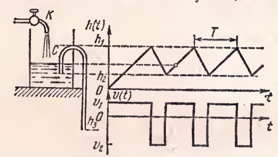


Рис. 7.1. Типичный пример релаксационных колебаний.

При наличии обратной связи в них устанавливаются самоподдерживающиеся почти синусоидальные колебания. Частота таких колебаний задается резонатором.

Но автоколебания могут происходить и в системах, не содержащих резонатора. Колебания при этом, как правило, не являются гармоническими. Типичными примерами таких

систем могут служить генератор пилообразных колебаний на неоновой лампе и гидравлическое устройство, показанное на рис. 7.1. В сосуд, снабженный сифоном C, с постоянной скоростью натекает вода из крана K. Пока сифон не заполнен водой, уровень воды в сосуде растет со временем по линейному закону. Но как только уровень достигает высоты h_1, сифон срабатывает и уровень воды в

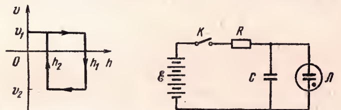


Рис. 7.2. Фазовая диаграмма релаксационных колебаний, показанных на рис. 7.1.

Рис. 7.3. Генератор пилообразных колебаний на неоновой лампе.

сосуде падает до значения h_2, после чего сосуд снова начинает заполняться водой из крана. Скорость опорожнения сосуда через сифон v_2 можно сделать гораздо больше скорости его наполнения через кран v_1, так как скорость воды в сифоне зависит от разностей уровней h_2 и h_3. Далее описанный процесс будет повторяться периодически. Зависимости уровня воды h и скорости его изменения v от времени показаны в правой части рис. 7.1. Видно, что колебания уровня воды и скорости не являются синусоидальными. Соответствующая этим колебаниям фазовая диаграмма приведена на рис. 7.2.

Аналогичные процессы происходят в генераторе пилообразных колебаний на неоновой лампе. Его электрическая схема показана на рис. 7.3. Неоновая лампа $Л$ обладает тем свойством, что ток через нее не проходит до тех пор, пока приложенное к лампе напряжение не достигнет определенной величины, называемой напряжением зажигания $U_з$. Если после возникновения тлеющего разряда в лампе напряжение на ней несколько уменьшить, то лампа будет продолжать гореть. Ток через лампу прекратится лишь тогда, когда напряжение будет уменьшено до определенного значения, называемого напряжением гашения $U_г$.

При замыкании ключа конденсатор C начинает медленно заряжаться через сопротивление R. Как только напряжение на конденсаторе достигнет значения, равного напряжению зажигания лампы $U_з$, в лампе возникает газовый разряд, и конденсатор начинает быстро разряжаться через лампу, так как сопротивление горящей неоновой лампы очень мало. Когда напряжение на конденсаторе уменьшится до значения напряжения гашения $U_г$,

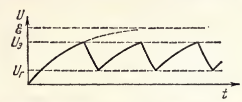


Рис. 7.4. Зависимость напряжения на конденсаторе от времени.

разряд в лампе прекращается, и конденсатор опять начинает заряжаться. Затем все повторяется снова. График зависимости напряжения на конденсаторе от времени $U(t)$ приведен на рис. 7.4.

Автоколебания, происходящие в генераторе на неоновой лампе и в рассмотренном выше гидравлическом устройстве, носят название релаксационных. Для таких колебаний характерно постоянное накопление энергии системой до некоторой величины, а затем быстрое «избавление» от накопленной энергии. Аналогом накопительного бачка в гидравлическом устройстве является конденсатор в генераторе пилообразного напряжения, аналогом сифона является неоновая лампа, а роль крана играет сопротивление R.

Возможные типы автоколебаний не исчерпываются рассмотренными примерами. Форма колебаний не обязательно бывает синусоидальной или пилообразной — она может быть какой угодно другой. Это относится не только к автоколебаниям, но и ко всем колебаниям вообще, включая и собственные, и вынужденные.

Синусоидальные колебания, конечно, занимают особое место. Во-первых, потому, что они очень часто встречаются: малые колебания вблизи устойчивого положения равновесия можно считать синусоидальными, так как в непо-

средственной окрестности минимума потенциальной энергии ее зависимость от смещения можно считать квадратичной. Во-вторых, такие колебания достаточно просты для математического исследования, и часто оказывается удобным представлять более сложные колебания как суперпозицию простых гармонических колебаний.

В общем случае несинусоидальные колебания сложнее синусоидальных. Но можно привести примеры таких

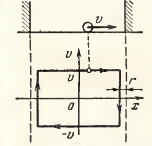


Рис. 7.5. Шарик, упруго отражающийся от стенок, и его фазовая траектория.

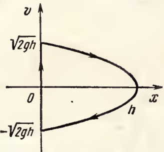


Рис. 7.6. Фазовая траектория шарика, отскакивающего от горизонтальной упругой плиты в поле тяжести.

систем, в которых несинусоидальные колебания могут быть легко исследованы. Приведем несколько примеров несинусоидальных собственных колебаний.

Простейший пример — это движущийся без трения по горизонтали упругий шарик, поочередно отражающийся от двух параллельных вертикальных стенок (рис. 7.5). Если удар о стенку считать абсолютно упругим, то неизменная по величине скорость шарика практически скачком меняет направление у стенки. Фазовая траектория движения шарика радиуса r показана на нижней части рис. 7.5. Обратим внимание, что скорость шарика, а следовательно, и период колебаний, в отличие от синусоидальных колебаний, зависят от начальных условий.

Другой пример — шарик, свободно падающий в поле тяжести и упруго отражающийся от горизонтальной плиты. Период колебаний шарика зависит от максимальной высоты подъема h, которая определяется начальными условиями, и равен $2\sqrt{2h/g}$. Зависимость скорости шарика от его

высоты x легко определяется с помощью закона сохранения энергии:

$$v(x) = \pm \sqrt{2g(h-x)}. \qquad (7.1)$$

Фазовая траектория, изображенная на рис. 7.6, состоит из вертикального участка, соответствующего мгновенному изменению скорости шарика при ударе о плиту, и параболы, определяемой уравнением (7.1). Для удобства начало отсчета высоты шарика x выбрано на высоте r над плитой, где r — радиус шарика.

ВОЛНЫ

§ 8. Колебания связанных маятников

Волны представляют собой процесс распространения колебаний. Волны бывают самой разнообразной природы: в зависимости от того, чтó колеблется и где распространяется, различают звуковые волны в упругой среде, волны

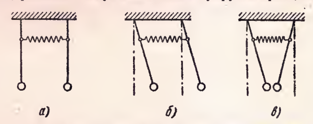


Рис. 8.1. Связанные маятники и их нормальные моды.

на поверхности воды, электромагнитные волны в пустоте или в веществе и т. п. Несмотря на различную физическую природу, волны любого типа, как и колебания разной природы, имеют очень много общего и подчиняются аналогичным закономерностям. Поэтому целесообразно изучать их совместно.

Чтобы получить представление о процессе распространения колебаний, рассмотрим простую механическую систему. Пусть два одинаковых маятника с жесткими стержнями соединены слабой пружиной ничтожной массы так, как показано на рис. 8.1, *а*. Маятники могут совершать

колебания в плоскости чертежа. Когда они висят вертикально, пружинка не деформирована.

В отличие от всех колебательных систем, рассмотренных выше, такие связанные маятники представляют собой механическую систему с двумя степенями свободы. Для задания механического состояния этой системы нужно указать положения и скорости обоих маятников. Какие колебания возможны в такой системе?

Рассмотрим сначала собственные колебания, которые возникают, если систему вывести из состояния равновесия и предоставить самой себе. Вид этих колебаний определяется начальными условиями. Если, например, отклонить оба маятника из положения равновесия в одну сторону на один и тот же угол и отпустить (рис. 8.1, *б*), то маятники будут качаться синхронно и пружинка будет все время оставаться в недеформированном состоянии. Это значит, что пружинка не оказывает на такое движение маятников никакого влияния. Каждый из маятников совершает гармоническое колебание, а амплитуды и фазы у них одинаковы. Частота этих колебаний ω_1 совпадает с частотой собственных гармонических колебаний одиночного маятника. Обозначая смещения маятников через x_1 и x_2, а начальное отклонение через A, можем написать

$$x_1(t) = A\cos\omega_1 t,\\ x_2(t) = A\cos\omega_1 t. \tag{8.1}$$

Но если отклонить маятники на одинаковый угол в противоположные стороны и отпустить (рис. 8.1, *в*), то маятники будут качаться в противофазе. Колебания маятников и в этом случае будут гармоническими, но частота колебаний ω_2 будет больше, чем ω_1. Действительно, при колебаниях такого типа в каждый момент времени деформированная пружина стремится повернуть маятники в ту же сторону, что и сила тяжести, т. е. в положение равновесия. Это приводит к увеличению частоты колебаний. Выражения для смещений маятников при начальном отклонении на величину A навстречу друг другу имеют вид

$$x_1(t) = A\cos\omega_2 t,\\ x_2(t) = -A\cos\omega_2 t. \tag{8.2}$$

Рассмотренные колебания, когда каждый из маятников совершает гармоническое движение, носят название нор-

мальных колебаний или нормальных мод. Число различных типов нормальных колебаний и соответствующих им частот равно числу степеней свободы системы. Поэтому в изучаемой системе двух связанных маятников других типов нормальных колебаний, кроме указанных выше, нет. Любое собственное колебание двух связанных маятников можно представить как суперпозицию их нормальных колебаний.

Рассмотрим, например, какое движение будут совершать маятники при суперпозиции нормальных колебаний (8.1) и (8.2) с одинаковыми амплитудами. Складывая между собой смещения каждого маятника при этих колебаниях, получим

$$\begin{aligned} x_1(t) &= A(\cos\omega_1 t + \cos\omega_2 t) = 2A\cos\frac{\Delta\omega}{2}t\cdot\cos\omega t, \\ x_2(t) &= A(\cos\omega_1 t - \cos\omega_2 t) = 2A\sin\frac{\Delta\omega}{2}t\cdot\sin\omega t, \end{aligned} \qquad (8.3)$$

где $\omega=\frac{1}{2}(\omega_1+\omega_2)$ — средняя частота, а $\Delta\omega=\omega_2-\omega_1$ — разность частот нормальных колебаний.

Если связывающая маятники пружинка очень слабая, то ее влияние на маятники мало, и поэтому частота ω_2 мало отличается от частоты ω_1 колебаний под действием только поля тяжести. В этом случае $\Delta\omega\ll\omega$ и правые части выражений (8.3) можно рассматривать как уравнения синусоидальных колебаний с частотой ω и медленно пульсирующими амплитудами. Графики этих колебаний показаны на рис. 8.2. Такое движение, получающееся в результате сложения гармонических колебаний с близкими частотами, носит название биений.

Из формул (8.3) видно, что в начальный момент $t=0$ $x_1=2A$, $x_2=0$. Это значит, что начальные условия, приводящие к такому движению, имеют вид: в начальный момент времени первый маятник смещен из положения равновесия вправо на величину $2A$, а второй маятник удерживается в положении равновесия. Легко видеть, что эти начальные условия представляют собой суперпозицию начальных условий для нормальных колебаний (8.1) и (8.2), как, разумеется, и должно быть.

Вся энергия колебаний в начальный момент сосредоточена у первого маятника. Но постепенно амплитуда его

колебаний, как видно из рис. 8.2, убывает, а амплитуда колебаний второго — возрастает. Через промежуток времени $T_0/2=\pi/\Delta\omega$, равный половине периода биений, вся энергия будет сосредоточена у второго маятника. Затем энергия колебаний снова переходит к первому маятнику и т. д. Процесс перекачки энергии колебаний от одного

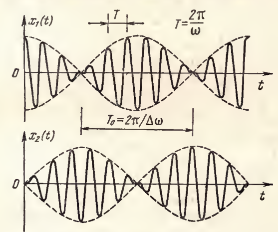


Рис. 8.2. Графики колебаний связанных маятников.

маятника к другому осуществляется благодаря связывающей маятники упругой пружине. Чем больше жесткость пружины, тем быстрее происходит передача энергии.

По сути дела, биения связанных маятников представляют собой волну в системе с двумя степенями свободы.

Опыт для наблюдения нормальных колебаний и биений связанных маятников можно осуществить очень просто. Для этого достаточно подвесить два одинаковых грузика на нитях одинаковой длины к горизонтальной нити на
50 некотором расстоянии друг от друга. Горизонтальная нить служит не только для подвешивания маятников, но и осуществляет связь между ними, т. е. играет роль еще и пружинки в описанном выше опыте. Эта связь будет эффективнее, если горизонтальную нить не натягивать сильно: она должна заметно провисать под тяжестью грузиков. Наблюдать колебания маятников удобнее в плоскостях, перпендикулярных горизонтальной нити. Если один из

маятников отклонить из положения равновесия и отпустить, то через некоторое время он остановится, но другой маятник будет раскачиваться. Затем снова раскачается первый маятник, а второй остановится и т. д.

Рассмотрим теперь установившиеся вынужденные колебания системы связанных маятников под действием синусоидальной внешней силы, считая трение пренебрежимо малым. Приложить эту силу можно таким же способом, как и при изучении вынужденных колебаний одиночного маятника. Можно взять длинную пружину с малой жесткостью и прикрепить ее, например, к первому маятнику, как показано на рис. 8.3. Другой конец этой длинной пружины следует привести в движение по закону $B\cos\omega t$. Действующая со стороны этой пружины на первый маятник сила будет синусоидальной, если амплитуда B левого конца пружины будет много больше амплитуды колебаний стержня первого маятника в точке закрепления пружины.

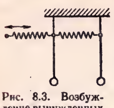


Рис. 8.3. Возбуждение вынужденных колебаний связанных маятников.

При установившихся вынужденных колебаниях каждый из маятников совершает гармоническое колебание с частотой, равной частоте внешнего воздействия. Если частота внешней силы очень мала, т. е. много меньше любой из частот нормальных колебаний, то оба маятника качаются с малыми амплитудами в фазе друг с другом и в фазе с вынуждающей силой. В предельном случае нулевой частоты, т. е. под действием постоянной внешней силы, маятники просто отклоняются в одну сторону в направлении действия силы и занимают новые положения равновесия.

По мере увеличения частоты внешней силы ω маятники по-прежнему колеблются в фазе, а амплитуды растут (рис. 8.4, *а*), пока частота ω не достигнет значения ω_1, т. е. частоты первого нормального колебания связанных маятников. При малом трении на этой частоте в системе отчетливо проявляется резонанс — амплитуда установившихся колебаний максимальна. В этом случае вынужденные колебания очень похожи на свободные нормальные колебания системы маятников на частоте ω_1 (рис. 8.1, *б*). Внешняя сила при этом лишь компенсирует затухание.

На рис. 8.4 направление действия вынуждающей силы показано стрелками.

Прн дальнейшем увеличенин частоты амплнтуды вынужденных установившнхся колебаний маятнников убывают, н при некоторой частоте ω_3, лежащей в промежутке между частотамн нормальных колебаний ω_1 и ω_2, амплнтуда колебаний первого маятнинка обращается в нуль. Этот маятник стонт неподвнжно, в то время как второй маятник качается с небольшой амплнтудой в протнвофазе с внешней снлой (рнс. 8.4, *б*, *в*). Первый маятник неподвижен, потому что действующне на него со стороны пружнн снлы в каждый момент времени уравновешнваются.

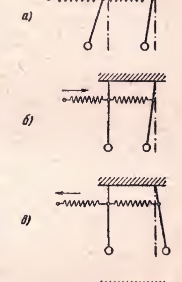


Рис. 8.4. Вынужденные колебания связанных маятников при разных значениях частоты вынуждающей силы.

Еслн частоту внешней силы еще немного увеличнть, то первый маятннк опять будет качаться, но уже в протнвофазе со вторым. Прн частоте $\omega=\omega_2$ в снстеме снова наступает резонанс. Амплнтуды резко возрастают, но, в отлнчие от резонанса на частоте ω_1, маятннкн качаются в противофазе друг с другом. Прн этом резонансе вынужденные колебания маятников очень похожи на свободные нормальные колебания на частоте ω_2 (рнс. 8.1, *в*).

Дальнейшее увелнчение частоты ω приводнт к тому, что амплнтуды вынужденных колебаннй маятнников снова убывают. Характер такнх колебаннй нллюстрируется рнс. 8.4, *г*.

Самым неожиданным в вынужденных колебаниях связанных маятников является, пожалуй, то, что на некоторой

частоте ω_3 тот маятник, на который действует внешняя сила, неподвижен. Глядя на рис. 8.4, *б* и *в*, легко сообразить, что частота ω_3, на которой это происходит, есть частота свободных колебаний второго маятника в этой системе, если первый закрепить неподвижно в положении равновесия. Строго говоря, амплитуда вынужденных колебаний первого маятника на частоте ω_3 равна нулю только при полном отсутствии трения. В противном случае первый маятник будет совершать колебания небольшой амплитуды, необходимые для поддержания стационарных колебаний второго маятника.

Это явление может быть использовано для борьбы с резонансом, ибо с его помощью интенсивные колебания можно перенести в дополнительно присоединенное устройство, колебания которого не причиняют вреда. Такой метод борьбы с вибрациями называется динамическим демпфированием и может применяться там, где частота вредных вибраций постоянна.

§ 9. Волны в упругих средах

В предыдущем параграфе было показано, что в системе связанных маятников колебательное движение не остается локализованным на одном маятнике: наличие связи между маятниками приводит к передаче колебательного движения от одного маятника к другому. Этот результат, полученный на конкретном примере, имеет общий характер — в системе связанных осцилляторов колебательное возбуждение будет распространяться. Распространение колебаний и представляет собой волновой процесс.

С давних пор наглядный образ волны всегда ассоциировался с волнами на поверхности воды. Но волны на воде представляют собой значительно более сложное явление, чем многие другие волновые процессы,— такие, как распространение звука или света в однородной изотропной среде. Поэтому естественно начинать изучение волнового движения не с волн на воде, а с более простых случаев.

Проще всего представить себе волну, распространяющуюся по бесконечной цепочке связанных маятников. С бесконечной цепочки мы начинаем для того, чтобы можно было рассматривать волну, распространяющуюся в одном

направлении, и не думать о возможном ее отражении от конца цепочки.

Если маятник, находящийся в начале цепочки, привести в гармоническое колебательное движение с некоторой частотой ω и амплитудой A, то колебательное движение будет распространяться по цепочке, и в отсутствие затухания любой другой маятник в цепочке будет повторять вынужденные колебания первого маятника с некоторым отставанием по фазе. Это запаздывание связано с тем, что распространение колебаний по цепочке происходит с некоторой конечной скоростью. Величина скорости распространения колебаний u зависит, как мы видели на примере двух связанных маятников, от жесткости соединяющей их пружинки, т. е. от того, насколько сильна связь между маятниками. Если первый маятник в цепочке движется по определенному закону, т. е. его смещение из положения равновесия есть заданная функция времени $x(t)$, то смещение маятника, отстоящего от начала цепочки на расстояние z, в любой момент времени t будет точно таким же, как смещение первого маятника в более ранний момент времени $t-z/u$, т. е. будет описываться функцией $x(t-z/u)$.

Пусть при гармонических колебаниях первого маятника его смещение из положения равновесия дается выражением

$$x(t)=A\cos\omega t. \qquad (9.1)$$

Каждый из маятников цепочки характеризуется тем расстоянием z, на которое он отстоит от начала цепочки; поэтому его смещение из положения равновесия при прохождении волны естественно обозначить через $x(t, z)$. Тогда, в соответствии со сказанным выше, имеем

$$x(t, z)=A\cos\omega\left(t-\frac{z}{u}\right). \qquad (9.2)$$

Описываемая уравнением (9.2) волна называется монохроматической. Характерным признаком монохроматической волны является то, что каждый из маятников совершает синусоидальное колебание определенной частоты.

Распространение волны по цепочке маятников сопровождается переносом энергии и импульса. Но никакого переноса массы при этом не происходит: каждый маятник, совершая колебания около положения равновесия, в среднем остается на месте.

В зависимости от того, в каком направлении происходят колебания маятников, говорят о волнах разной поляризации. Если колебания маятников происходят вдоль направления распространения волны, то волна называется продольной, если поперек — то поперечной. Обычно волны разной поляризации распространяются с разными скоростями.

Рассмотренная цепочка связанных маятников представляет собой пример механической системы с сосредоточенными параметрами. Другой пример системы с сосредоточенными параметрами, в которой могут распространяться

Рис. 9.1. Цепочка шариков, соединенных пружинками,— пример системы с сосредоточенными параметрами.

волны,— это цепочка шариков, связанных легкими пружинками (рис. 9.1). В такой системе инертные свойства сосредоточены у шариков, а упругие — у пружинок. При распространении волны кинетическая энергия колебаний локализована на шариках, а потенциальная — на пружинках.

Легко сообразить, что такую цепочку соединенных пружинками шариков можно рассматривать как модель одномерной системы с распределенными параметрами, например упругой струны. В струне каждый элемент длины обладает одновременно и массой, т. е. инертными свойствами, и жесткостью, т. е. упругими свойствами.

Рассмотрим поперечную монохроматическую волну, распространяющуюся в бесконечной натянутой струне. Предварительное натяжение струны необходимо потому, что ненатянутая гибкая струна, в отличие от твердого стержня, обладает упругостью только по отношению к деформации растяжения, но не сжатия.

Монохроматическая волна в струне описывается тем же выражением (9.2), что и волна в цепочке маятников. Однако теперь роль отдельного маятника играет каждый элемент струны, поэтому переменная z в уравнении (9.2), характеризующая равновесное положение маятника, принимает непрерывные значения. Смещение любого элемента струны из равновесного положения при прохождении волны $x(t, z)$

есть функция двух переменных: времени t и равновесного положения этого элемента z. Если в формуле (9.2) зафиксировать z, т. е. рассматривать определенный элемент струны, то функция $x(t, z)$ при фиксированном z дает смещение выделенного элемента струны в зависимости от времени. Это смещение представляет собой гармоническое колебание с частотой ω и амплитудой A:

$$x(t, z)=A\cos(\omega t+\alpha). \tag{9.3}$$

Начальная фаза колебаний этого элемента струны α равна $-\frac{\omega}{u}z$, т. е. зависит от его равновесного положения. Все элементы струны при прохождении монохроматической

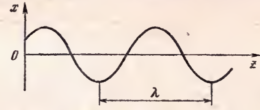


Рис. 9.2. Картина смещений разных точек струны в один и тот же момент времени.

волны совершают гармонические колебания одинаковой частоты и амплитуды, но различающиеся по фазе.

Если в формуле (9.2) зафиксировать t, т. е. рассматривать всю струну в один и тот же момент времени, то функция $x(t, z)$ при фиксированном t дает мгновенную картину смещений всех элементов струны — как бы моментальную фотографию волны. На этой «фотографии» мы увидим застывшую синусоиду (рис. 9.2). Период этой синусоиды, т. е. расстояние между соседними горбами или впадинами, называется длиной волны λ. Из формулы (9.2) можно найти, что длина волны связана с частотой ω и скоростью волны u соотношением

$$\lambda=\frac{2\pi}{\omega}u=uT, \tag{9.4}$$

где T — период колебаний. Картину распространения волны можно представить себе, если эту «застывшую» синусоиду привести в движение вдоль оси z со скоростью u. Две последовательные «моментальные фотографии» волны в мо-

менты времени t и $t+\Delta t$ показаны на рис. 9.3. Видно, что длина волны λ равна расстоянию, проходимому любым горбом за период колебаний T, в соответствии с формулой (9.4).

Определим скорость распространения монохроматической поперечной волны в струне. Будем считать, что амплитуда A мала по сравнению с длиной волны: $A \ll \lambda$. Пусть

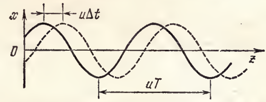


Рис. 9.3. Распространение волны можно представить себе как движение вдоль оси z синусоиды, показанной на рис. 9.2.

волна бежит направо со скоростью u. Перейдем в новую систему отсчета, движущуюся вдоль струны со скоростью, равной скорости волны u. Эта система отсчета также является инерциальной, и, следовательно, в ней справедливы законы Ньютона. Из этой системы отсчета волна кажется застывшей синусоидой, а вещество струны скользит вдоль этой синусоиды налево: любой предварительно окрашенный элемент струны будет казаться убегающим вдоль синусоиды налево со скоростью u.

Рассмотрим в этой системе отсчета элемент струны длиной Δl, которая много меньше длины волны λ, в тот момент, когда он находится на гребне синусоиды (рис. 9.4). Применим к этому элементу второй закон Ньютона. Силы, действующие на элемент со стороны соседних участков струны, показаны в выделенном кружке на рис. 9.4. Поскольку рассматривается поперечная волна, в которой смещения элементов струны перпендикулярны направлению распространения волны, то горизонтальная составляющая силы натяжения F постоянна вдоль всей струны. Так как длина рассматриваемого участка $\Delta l \ll \lambda$, то направления сил натяжения, действующих на выделенный элемент, почти горизонтальны, а их величину можно считать равной F. Равнодействующая этих сил направлена вниз и равна $F\,\Delta\alpha$. Скорость рассматриваемого элемента равна u и направлена влево,

а малый участок его синусоидальной траектории вблизи горба можно считать дугой окружности радиуса R. Поэтому ускорение этого элемента струны направлено вниз и равно u^2/R. Массу элемента струны можно представить в виде $\rho S \Delta l$, где ρ — плотность материала струны, а S — площадь сечения, которые ввиду малости деформаций при распространении волны можно считать постоянными. На основании второго закона Ньютона

$$F \Delta\alpha = \rho S \Delta l \frac{u^2}{R}. \quad (9.5)$$

Учитывая, что $\Delta l = R \Delta\alpha$ (см. рис. 9.4), получим

$$u = \sqrt{\frac{F}{\rho S}}. \quad (9.6)$$

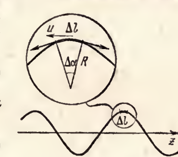


Рис. 9.4. К выводу выражения для скорости поперечной волны в натянутой струне.

Это и есть искомая скорость распространения поперечной монохроматической волны малой амплитуды в натянутой струне. Видно, что она зависит только от механического напряжения натянутой струны F/S и от ее плотности ρ и не зависит от амплитуды и длины волны. Это значит, что поперечные волны любой длины распространяются в натянутой струне с одинаковой скоростью. Если в струне одновременно распространяются, например, две волны с одинаковыми амплитудами и близкими частотами ω_1 и ω_2, то вследствие равенства их скоростей уравнение результирующей волны будет иметь вид

$$x(t, z) = A\cos\omega_1\left(t - \frac{z}{u}\right) + A\cos\omega_2\left(t - \frac{z}{u}\right) = \\ = 2A\cos\frac{\Delta\omega}{2}\left(t - \frac{z}{u}\right)\cdot\cos\omega\left(t - \frac{z}{u}\right), \quad (9.7)$$

где $\omega = \frac{1}{2}(\omega_1 + \omega_2)$, а $\Delta\omega = \omega_1 - \omega_2$. «Моментальные фотографии» монохроматических волн и результирующей волны показаны на рис. 9.5. Там, где горб одной волны совпадает с горбом другой, в результирующей волне смещение максимально. Поскольку соответствующие отдельным волнам

синусоиды бегут вдоль оси z с одинаковой скоростью u, то и результирующая кривая бежит с той же самой скоростью, не меняя своей формы. Оказывается, что это справедливо для волнового возмущения любой формы: поперечные волны произвольного вида распространяются в натянутой струне, не меняя своей формы.

Если скорость распространения монохроматических волн не зависит от длины волны или частоты, то говорят, что отсутствует дисперсия. Сохранение формы любой волны при ее распространении есть следствие отсутствия дисперсии.

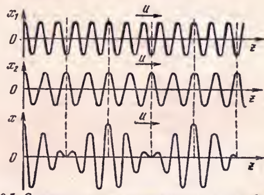


Рис. 9.5. Сложение двух монохроматических волн с близкими частотами.

Дисперсия отсутствует для волн любого вида, распространяющихся в сплошных упругих средах. Это обстоятельство позволяет очень легко найти скорость продольных волн.

Рассмотрим, например, длинный упругий стержень площади S, в котором распространяется продольное возмущение с крутым передним фронтом. Пусть в некоторый момент времени t фронт волны, перемещаясь со скоростью u, дошел до точки с координатой z; справа от фронта все точки стержня еще покоятся. Спустя промежуток времени Δt фронт переместится вправо на расстояние $u\,\Delta t$ (рис. 9.6). В пределах этого слоя все частицы движутся с одной и той же скоростью v. Спустя промежуток времени Δt частицы стержня, находившиеся в момент t на фронте волны, переместятся вдоль стержня на расстояние $v\,\Delta t$.

Применим к вовлеченной за время Δt в волновой процесс массе стержня $\Delta m = \rho S u \Delta t$ закон сохранения импульса:

$$v \Delta m = v\rho S u \Delta t = F \Delta t. \qquad (9.8)$$

Величину действующей на массу Δm силы F выразим через деформацию элемента стержня с помощью закона Гука:

$$\frac{\Delta l}{l} = \frac{F}{ES}. \qquad (9.9)$$

Длина l выделенного элемента стержня равна $u \Delta t$, а изменение его длины Δl под действием силы F равно $v \Delta t$. Поэтому с помощью (9.9) находим

$$F = ES\frac{v}{u}. \qquad (9.10)$$

Подставляя это значение в (9.8), получаем

$$u = \sqrt{\frac{E}{\rho}}. \qquad (9.11)$$

Скорость продольных звуковых волн в упругом стержне зависит только от модуля Юнга E и плотности ρ. Легко убедиться, что в большинстве металлов эта скорость составляет примерно 5 км/с. Скорость продольных волн в упругой среде всегда больше скорости поперечных. Сравним, например, скорости продольных и поперечных волн u_l и u_t в натянутой гибкой струне. Поскольку при малых деформациях упругие постоянные не зависят от величины приложенных сил, то скорость продольных волн в натянутой струне не зависит от величины ее предварительного натяжения и определяется формулой (9.11). Для того чтобы сравнить эту скорость с найденной ранее скоростью поперечных волн u_t, выразим натягивающую струну силу F, входящую в формулу (9.6), через относительную деформацию струны $\varepsilon = \Delta l / l_0$, обусловленную этим предварительным натяжением: $\varepsilon = F/(SE)$. Подставляя значение F в

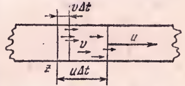


Рис. 9.6. К выводу выражения для скорости продольных волн в стержне.

формулу (9.6), получаем

$$u_t = \sqrt{\frac{\varepsilon E}{\rho}}. \qquad (9.12)$$

Таким образом, скорость поперечных волн в натянутой струне u_t оказывается значительно меньше скорости продольных волн, так как относительное растяжение струны ε много меньше единицы.

§ 10. Энергия волн

При распространении волн происходит передача энергии без переноса вещества. Энергия волны в упругой среде состоит из кинетической энергии совершающих колебания частиц вещества и из потенциальной энергии упругой деформации среды.

Рассмотрим, например, продольную волну в упругом стержне. В фиксированный момент времени кинетическая энергия распределена по объему стержня неравномерно, так как одни точки стержня в этот момент покоятся, другие, напротив, движутся с максимальной скоростью. То же самое справедливо и для потенциальной энергии, так как в этот момент какие-то элементы стержня не деформированы, другие же деформированы максимально. Поэтому при рассмотрении энергии волны естественно вводить плотность кинетической и потенциальной энергий. Плотность энергии волны в каждой точке среды не остается постоянной, а периодически меняется при прохождении волны: энергия распространяется вместе с волной.

Начнем с плотности кинетической энергии в монохроматической упругой волне, описываемой уравнением (9.2):

$$x(t, z) = A\cos\omega\left(t - \frac{z}{u}\right). \qquad (10.1)$$

Выделим в стержне малый элемент длины между плоскостями z и $z+\Delta z$ такой, что его длина Δz в недеформированном состоянии много меньше длины волны λ. Тогда скорости v всех частиц стержня в этом элементе при распространении волны можно считать одинаковыми. С помощью формулы (10.1) находим скорость $v=\dot{x}$, рассматривая $x(t, z)$ как функцию времени и считая величину z, характеризующую положение рассматриваемого элемента стержня,

фиксированной:

$$v(t,\ z)=\dot{x}=-\omega A \sin\omega\left(t-\frac{z}{u}\right). \qquad (10.2)$$

Масса выделенного элемента стержня Δm равна $\rho S\,\Delta z$, поэтому его кинетическая энергия $\Delta E_к$ в момент времени t есть

$$\Delta E_к=\frac{1}{2}\Delta m v^2=\frac{1}{2}\rho S\,\Delta z\,\omega^2 A^2 \sin^2\omega\left(t-\frac{z}{u}\right). \qquad (10.3)$$

С помощью выражения (10.3) находим плотность кинетической энергии $w_к(t,\ z)$ в точке z в момент времени t:

$$w_к(t,\ z)=\frac{\Delta E_к}{S\,\Delta z}=\frac{1}{2}\rho\omega^2 A^2 \sin^2\omega\left(t-\frac{z}{u}\right). \qquad (10.4)$$

Перейдем к вычислению плотности потенциальной энергии волны. Поскольку длина выделенного элемента стержня мала по сравнению с длиной волны, то вызываемую волной деформацию этого элемента можно считать однородной. Поэтому потенциальную энергию деформации $\Delta E_п$, в соответствии с формулой (8.9) раздела «Механика», можно записать в виде

$$\Delta E_п=\frac{1}{2}S\,\Delta z\,E\left(\frac{\Delta l}{\Delta z}\right)^2, \qquad (10.5)$$

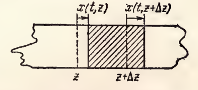


Рис. 10.1. К выводу формулы для относительного удлинения выделенного элемента стержня.

где Δl — удлинение рассматриваемого элемента стержня Δz, вызванное проходящей волной. Для нахождения этого удлинения нужно рассмотреть положение плоскостей, ограничивающих выделенный элемент, в некоторый момент времени t. Мгновенное положение любой плоскости, равновесное положение которой характеризуется координатой z, определяется функцией $x(t,\ z)$, рассматриваемой как функция z при фиксированном t. Поэтому удлинение Δl рассматриваемого элемента стержня, как видно из рис. 10.1, равно

$$\Delta l=x(t,\ z+\Delta z)-x(t,\ z).$$

Относительное удлинение этого элемента есть

$$\frac{\Delta l}{\Delta z}=\frac{x(t,\ z+\Delta z)-x(t,\ z)}{\Delta z}.$$

Если в этом выражении перейти к пределу при $\Delta z \to 0$, то оно превращается в производную функции $x(t, z)$ по переменной z при фиксированном t. С помощью формулы (10.1) получаем

$$\frac{\Delta l}{\Delta z} \to \frac{\omega}{u} A \sin \omega \left(t - \frac{z}{u}\right). \qquad (10.6)$$

Теперь выражение для потенциальной энергии (10.5) принимает вид

$$\Delta E_п = \frac{1}{2} S \Delta z E \left(\frac{\omega}{u} A\right)^2 \sin^2 \omega \left(t - \frac{z}{u}\right), \qquad (10.7)$$

а плотность потенциальной энергии $w_п(t, z)$ в точке z в момент времени t есть

$$w_п(t, z) = \frac{\Delta E_п}{S \Delta z} = \frac{1}{2} E \frac{\omega^2}{u^2} A^2 \sin^2 \omega \left(t - \frac{z}{u}\right). \qquad (10.8)$$

Поскольку скорость распространения продольных волн $u = \sqrt{E/\rho}$, то правые части в формулах (10.8) и (10.4) совпадают. Это значит, что в бегущей продольной упругой волне плотности кинетической и потенциальной энергий равны в любой момент времени в любой точке среды. Зависимость плотности энергии волны $w = w_к + w_п$ от координаты z в фиксированный момент времени t показана на рис. 10.2.

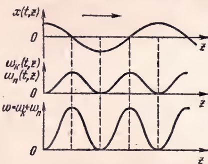


Рис. 10.2. Смещение частиц среды и плотность энергии в бегущей волне для фиксированного момента времени.

Обратим внимание на то, что, в отличие от локализованных колебаний (осциллятор), где кинетическая и потенциальная энергии изменяются в противофазе (рис. 1.4), в бегущей волне колебания кинетической и потенциальной энергий происходят в одинаковой фазе. Кинетическая и потенциальная энергии в каждой точке среды одновременно достигают максимальных значений и одновременно обращаются в нуль.

Равенство мгновенных значений плотностей кинетической и потенциальной энергий есть общее свойство бегущих волн, т. е. волн, распространяющихся в определенном направлении. Легко убедиться, что это справедливо, например, и для поперечных волн в натянутой гибкой струне.

Смещение элементов струны при прохождении поперечной волны происходит перпендикулярно направлению распространения волны, но описывается той же самой формулой (10.1). Поэтому кинетическая энергия и ее плотность в бегущей поперечной волне даются теми же формулами (10.3), (10.4), что и в продольной волне.

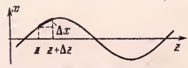


Рис. 10.3. К вычислению вызванного волной удлинения элемента струны.

При вычислении потенциальной энергии волны учтем, что предварительно натянутая струна уже деформирована и обладает некоторой потенциальной энергией, не имеющей никакого отношения к энергии волны. Нас интересует только та часть потенциальной энергии деформированной струны, которая обусловлена дополнительным растяжением струны, вызываемым волной. Для поперечной волны малой амплитуды сила натяжения струны в любом месте мало отличается от ее значения F в отсутствие волны. Поэтому потенциальная энергия любого элемента струны Δz, связанная с волной, равна произведению силы натяжения струны F на малое удлинение этого элемента Δl при прохождении волны. Это удлинение легко подсчитать с помощью теоремы Пифагора (рис. 10.3):

$$\Delta l = \sqrt{(\Delta z)^2 + (\Delta x)^2} - \Delta z.$$

При малых амплитудах Δx мало по сравнению с Δz, поэтому

$$\sqrt{(\Delta z)^2 + (\Delta x)^2} = \Delta z \sqrt{1 + \left(\frac{\Delta x}{\Delta z}\right)^2} \approx \Delta z \left[1 + \frac{1}{2}\left(\frac{\Delta x}{\Delta z}\right)^2\right].$$

Теперь выражение для удлинения Δl принимает вид

$$\Delta l \approx \frac{1}{2}\left(\frac{\Delta x}{\Delta z}\right)^2 \Delta z. \qquad (10.9)$$

При $\Delta z \to 0$ отношение $\Delta x/\Delta z$ превращается в производную функции $x(t, z)$ по z при фиксированном t. Поэтому

$$\Delta l \longrightarrow \Delta z \cdot \frac{1}{2} \left(\frac{\omega A}{u}\right)^2 \sin^2 \omega \left(t - \frac{z}{u}\right) \qquad (10.10)$$

и связанная с волной потенциальная энергия $\Delta E_п$ участка струны Δz принимает вид

$$\Delta E_п = F \, \Delta l = \frac{1}{2} F \frac{\omega^2}{u^2} A^2 \Delta z \cdot \sin^2 \omega \left(t - \frac{z}{u}\right). \quad (10.11)$$

Учитывая, что скорость поперечной волны $u = \sqrt{F/(S\rho)}$, и вычисляя плотность потенциальной энергии $w_п = \Delta E_п/(S \, \Delta z)$, убеждаемся, что она равна плотности кинетической энергии $w_к$.

Энергия бегущей волны не остается локализованной: она перемещается вместе с волной со скоростью u. Имея выражение для объемной плотности энергии волны w:

$$w(t, z) = w_к + w_п = \rho\omega^2 A^2 \sin^2 \omega \left(t - \frac{z}{u}\right), \quad (10.12)$$

легко найти поток энергии $\Delta\Phi$, переносимый волной за единицу времени через произвольную площадку ΔS, перпендикулярную направлению распространения волны:

$$\Delta\Phi = wu \, \Delta S. \qquad (10.13)$$

Величина

$$j = wu \qquad (10.14)$$

носит название плотности потока энергии волны. Она была впервые введена русским физиком Н. А. Умовым и имеет смысл энергии, переносимой волной в единицу времени через единичную площадку, перпендикулярную направлению распространения волны.

Плотность потока энергии, как видно из выражения (10.12), осциллирует с удвоенной частотой 2ω. Среднее по времени значение плотности энергии в любой точке, через которую проходит волна, равно $\frac{1}{2}\rho\omega^2 A^2$. Поэтому перенос энергии волной можно характеризовать средним значением плотности потока энергии $\langle j \rangle$:

$$\langle j \rangle = \frac{1}{2} \rho\omega^2 A^2 u. \qquad (10.15)$$

Отметим, что средний поток энергии волны пропорционален квадрату амплитуды и квадрату частоты.

До сих пор мы рассматривали волны, распространяющиеся в системе, имеющей бесконечную протяженность только по одному направлению: в цепочке маятников, в струне, в стержне. Но волны могут распространяться и в среде, имеющей бесконечные размеры по всем направлениям. В такой сплошной среде волны бывают разного вида в зависимости от способа их возбуждения.

Если, например, волна возникает в результате гармонических колебаний бесконечной плоскости, то в однородной среде она распространяется в направлении, перпендикулярном к этой плоскости. В такой волне смещение всех точек среды, лежащих на любой плоскости, перпендикулярной направлению распространения, происходит совершенно одинаково. Если в среде не происходит поглощения энергии волны, то амплитуда колебаний точек среды всюду одинакова и их смещение дается формулой (10.1). Такая волна называется плоской.

Волну другого вида — сферическую — создает в однородной изотропной упругой среде пульсирующий шар. Такая волна распространяется с одинаковой скоростью по всем направлениям. Ее волновые поверхности, т. е. поверхности постоянной фазы, представляют собой концентрические сферы. В отсутствие поглощения энергии в среде легко определить зависимость амплитуды сферической волны от расстояния до центра. Поскольку поток энергии волны, пропорциональный, согласно (10.15), квадрату амплитуды, одинаков через любую сферу, то амплитуда волны убывает обратно пропорционально расстоянию r от центра: $A \sim 1/r$. Уравнение продольной сферической волны имеет вид

$$x(t,\ r) = a\frac{r_0}{r}\cos\omega\left(t - \frac{r}{u}\right), \qquad (10.16)$$

где a есть амплитуда колебаний на расстоянии r_0 от центра волны.

§ 11. Интерференция волн. Стоячие волны

При одновременном распространении нескольких волн смещения частиц среды представляют собой векторную сумму смещений, которые имели бы место при распространении каждой волны в отдельности. Иначе говоря, волны просто

накладываются одна на другую, не искажая друг друга. Этот экспериментальный факт был известен еще Леонардо да Винчи, который заметил, что круги волн на воде от разных источников проходят друг сквозь друга и распространяются дальше, не претерпев никаких изменений. Утверждение о независимом распространении нескольких волн носит название принципа суперпозиции для волнового движения.

Мы уже рассматривали распространение в одном направлении двух одинаково поляризованных монохроматических волн с близкими частотами. В результате наложения таких волн получается почти синусоидальная волна с периодически меняющейся в пространстве амплитудой. «Моментальная фотография» такой волны выглядит как следующие друг за другом группы волн (рис. 9.5), а вызываемое волной колебание в какой-либо фиксированной точке имеет характер биений (рис. 8.2).

Особый интерес представляет случай сложения так называемых когерентных волн, т. е. волн от согласованных источников. Простейшим примером когерентных волн являются монохроматические волны одинаковой частоты с постоянной разностью фаз. Для истинно монохроматических волн требование постоянной разности фаз будет лишним, так как они являются бесконечно протяженными в пространстве и во времени и две такие волны одинаковой частоты всегда имеют постоянную разность фаз. Но реальные волновые процессы, даже близкие к монохроматическим, всегда имеют конечную протяженность. Для того чтобы такие квазимонохроматические волны, представляющие собой последовательности отдельных синусоидальных цугов, были когерентными, требование постоянной разности фаз является обязательным. Строго говоря, понятие когерентности волн является более сложным, чем описано выше. Подробнее мы познакомимся с ним в разделе «Оптика».

При сложении когерентных волн наблюдаются явления интерференции, заключающиеся в том, что вызываемая этими волнами картина колебаний является стационарной, т. е. в каждой точке происходят колебания с не зависящей от времени амплитудой. Разумеется, в разных точках амплитуды колебаний будут различаться.

Для расчета интерференционных картин нужно уметь складывать колебания одинаковой частоты, происходящие

в одном направлении. Проще всего найти результат сложения таких колебаний, воспользовавшись представлением гармонических колебаний с помощью векторных диаграмм. Пусть складываемые колебания имеют вид

$$x_1(t) = A_1 \cos(\omega t + \alpha_1), \quad x_2(t) = A_2 \cos(\omega t + \alpha_2). \quad (11.1)$$

Этим колебаниям можно сопоставить векторы A_1 и A_2 на диаграмме, изображенной на рис. 11.1. Поскольку колебания имеют одинаковую частоту ω, сопоставляемые им векторы вращаются с одинаковой угловой скоростью, так что их взаимное расположение остается неизменным. Результирующему колебанию будет соответствовать вектор A, равный сумме векторов A_1 и A_2, и это колебание будет иметь вид

$$x(t) = x_1(t) + x_2(t) = A \cos(\omega t + \alpha). \quad (11.2)$$

Рис. 11.1. Сложение гармонических колебаний одинаковой частоты с помощью векторной диаграммы.

Амплитуду A и начальную фазу α результирующего колебания легко найти с помощью рис. 11.1. Применяя теорему косинусов, имеем

$$A^2 = A_1^2 + A_2^2 + 2A_1A_2 \cos(\alpha_1 - \alpha_2), \quad (11.3)$$

а для $\operatorname{tg}\alpha$ справедливо выражение

$$\operatorname{tg}\alpha = \frac{A_1 \sin\alpha_1 + A_2 \sin\alpha_2}{A_1 \cos\alpha_1 + A_2 \cos\alpha_2}.$$

Поскольку энергия колебаний пропорциональна квадрату амплитуды, то из выражения (11.3) видно, что энергия результирующего колебания в общем случае не равна сумме энергий складываемых колебаний. Например, если амплитуды A_1 и A_2 одинаковы, амплитуда результирующего колебания может иметь значения от $A = 0$ при $\alpha_1 - \alpha_2 = \pi$ (складываемые колебания происходят в противофазе) до $A = 2A_1$ при $\alpha_1 - \alpha_2 = 0$. В первом случае энергия результирующего колебания равна нулю, а во втором она в два раза больше суммы энергий складываемых колебаний.

Полученные результаты позволяют легко находить картину интерференции двух когерентных волн одинаковой по-

ляризации. Пусть, например, два когерентных источника, находящиеся на расстоянии d друг от друга, создают сферические волны, интерференция которых наблюдается в точке P (рис. 11.2). Если расстояния l_1 и l_2 от источников до точки наблюдения P велики по сравнению с расстоянием d между источниками, то амплитуды обеих волн в точке наблюдения будут практически одинаковыми. Одинаковыми будут и направления смещений точек среды, вызываемых этими волнами в месте наблюдения. Результат интерференции в точке P будет зависеть от разности фаз между волнами, приходящими в эту точку. Если источники совершают колебания в одинаковой фазе, то разность фаз волн в точке P зависит только от разности хода l волн от источников до точки наблюдения: $l=l_2-l_1$. Если эта разность хода равна целому числу длин волн, $l=k\lambda$, то волны приходят в точку P в фазе и, складываясь, дают колебание с удвоенной амплитудой. Если же разность хода равна нечетному числу полуволн, $l=(2k+1)\lambda/2$, то волны приходят в точку P в противофазе и «гасят» друг друга: амплитуда результирующего колебания равна нулю. При промежуточных значениях разности хода амплитуда колебаний в точке наблюдения, как следует из формулы (11.3), принимает определенное значение в промежутке между указанными предельными случаями. Каждая точка среды характеризуется определенным значением амплитуды колебаний, которое не меняется со временем. Распределение этих амплитуд в пространстве называется интерференционной картиной.

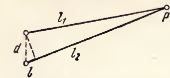


Рис. 11.2. К интерференции волн от двух точечных источников.

Гашение колебаний в одних местах и усиление в других при интерференции волн не связаны, вообще говоря, с какими-либо превращениями энергии колебаний. В точках, где колебания от двух волн гасят друг друга, энергия волн отнюдь не превращается в другие виды, например в тепло. Все сводится лишь к перераспределению потока энергии в пространстве, так что минимумы энергии колебаний в одних местах компенсируются максимумами в других в полном соответствии с законом сохранения энергии.

Однако в явлениях интерференции бывают случаи, когда на первый взгляд кажется, что закон сохранения энергии нарушается. Один из таких кажущихся парадоксов возникает в случае, когда расстояние между двумя одинаковыми когерентными источниками монохроматических волн значительно меньше длины волны: $d \ll \lambda$. Если источники совершают колебания в одинаковой фазе, то в любую точку во́лны от этих источников приходят почти в фазе, так как разность хода волн l много меньше длины волны λ. Волны всюду усиливают друг друга, нигде не происходит гашения волн. Поэтому амплитуда колебаний в любом месте оказывается почти вдвое больше амплитуды колебаний, вызываемых волной от каждого источника, а энергия — в четыре раза больше. Полный поток энергии оказывается почти вдвое больше суммы потоков, создаваемых каждым источником в отдельности. И тем не менее никакого противоречия с законом сохранения энергии здесь нет. Оказывается, что каждый источник при той же амплитуде колебаний действительно излучает вдвое больше энергии, когда рядом с ним находится другой такой же источник, колеблющийся в фазе с ним. Так происходит потому, что источники взаимодействуют через создаваемые ими волны. Работа, которая совершается при приведении источников волн в действие, будет в этом случае вдвое больше. Сопротивление движению каждого источника будет обусловлено не только той волной, которую он излучает сам, но и волной, излучаемой вторым источником. В результате при прежней амплитуде колебаний каждый источник развивает вдвое бо́льшую мощность.

Наблюдать описанные явления особенно удобно не для упругих механических волн, а для электромагнитных волн, излучаемых расположенными рядом антеннами. Если в этих антеннах колебания происходят синфазно, то каждый передатчик, питающий отдельную антенну, развивает вдвое бо́льшую мощность.

Аналогичный парадокс возникает и в случае, когда расположенные рядом когерентные источники совершают колебания в противофазе. Излучаемые такими источниками волны, интерферируя, всюду почти полностью гасят друг друга, так что амплитуда результирующей волны везде близка к нулю. Взаимодействие источников через излучаемые ими волны в этом случае приводит к тому, что энергия

вообще почти не излучается. Фактически образуется почти замкнутая система, в которой энергия только переходит из одной антенны к другой и обратно. Несмотря на значительную амплитуду колебаний в антеннах, развиваемая передатчиками мощность практически равна нулю: передатчики только восполняют потери на активных сопротивлениях антенн.

Очень разнообразны и многогранны проявления интерференции волн в оптике. В соответствующем разделе будет рассмотрен целый ряд интересных примеров интерференции.

Для наблюдения устойчивой интерференционной картины не обязательно иметь два независимых когерентных источника. Вторую, когерентную с исходной волну можно получить в результате отражения исходной волны от границы среды, в которой происходит распространение волн. В этом случае интерферируют падающая и отраженная волны.

Если плоская монохроматическая волна падает по нормали на плоскую границу раздела двух сред, то в результате отражения от границы возникает также плоская волна, распространяющаяся в обратном направлении. Аналогичное явление происходит при отражении распространяющейся в струне волны от закрепленного или свободного конца струны. При равенстве амплитуд падающей и отраженной волн в результате интерференции образуется стоячая волна. В стоячей волне, как и вообще при интерференции волн, каждая точка среды совершает гармоническое колебание с некоторой амплитудой, которая, в отличие от случая бегущей волны, в разных точках среды имеет разные значения.

Найдем распределение амплитуд в стоячей волне. Для определенности будем рассматривать поперечную волну в гибкой струне. Пусть уравнение падающей волны имеет вид

$$x_1(t,\ z) = A\cos\omega\left(t-\frac{z}{u}\right). \qquad (11.4)$$

Это уравнение описывает волну, бегущую со скоростью u в положительном направлении оси z из $-\infty$ в ∞. Предположим, что в точке $z=0$ волна испытывает отражение. Этого можно добиться, либо закрепив струну в этой точке, либо перерезав ее. В первом случае происходит отражение

волны от закрепленного конца струны, во втором — от свободного. Физические условия отражения в этих случаях разные, но амплитуда отраженной волны в обоих случаях равна амплитуде падающей, так как энергия волны не может передаваться далее. Если обозначить изменение фазы волны при отражении через δ, то уравнение отраженной волны записывается в виде

$$x_2(t,\ z) = A\cos\left[\omega\left(t + \frac{z}{u}\right) + \delta\right]. \qquad (11.5)$$

Оно описывает волну, распространяющуюся в отрицательном направлении оси z, и поэтому отличается от уравнения (11.4) заменой u на $-u$.

Как изменяется фаза волны при отражении? Колебание любой точки струны есть результат сложения гармонических колебаний, вызываемых падающей и отраженной волнами. Поэтому, если конец струны закреплен, то складываемые колебания в этой точке в любой момент времени должны погасить друг друга, т. е. они должны происходить в противофазе: для закрепленного конца струны $\delta = \pi$. Напротив, если конец струны в точке $z=0$ свободен, то амплитуда результирующего колебания в этой точке должна быть максимальной, т. е. падающая и отраженная волны имеют здесь одинаковую фазу: $\delta = 0$. Теперь уравнение отраженной волны (11.5) можно записать в виде

$$x_2(t,\ z) = \pm A\cos\omega\left(t + \frac{z}{u}\right). \qquad (11.6)$$

Знак «+» соответствует отражению от свободного конца, а знак «—» — отражению от закрепленного конца. Для определенности рассмотрим случай, когда конец струны в точке $z=0$ закреплен. Сложение падающей и отраженной волн приводит к следующему результату:

$$x(t,\ z) = x_1 + x_2 = A\left[\cos\omega\left(t - \frac{z}{u}\right) - \cos\omega\left(t + \frac{z}{u}\right)\right] = 2A\sin\frac{\omega z}{u}\sin\omega t. \qquad (11.7)$$

Формула (11.7) показывает, что каждая точка струны совершает гармоническое колебание с частотой ω. Амплитуда для разных точек различна и зависит от координаты точки

по следующему закону:

$$A(z) = 2A\left|\sin\frac{\omega z}{u}\right|. \qquad (11.8)$$

Точки, в которых амплитуда колебаний струны максимальна, называются пучностями стоячей волны. Точки, в которых амплитуда колебаний равна нулю, называются узлами. Как видно из формулы (11.8), узлы находятся в точках, определяемых условием

$$\frac{\omega z}{u} = n\pi,$$

где $n = 0, -1, -2, \ldots$ Расстояние между соседними узлами равно $u\pi/\omega$. Так как π/ω равно половине периода колебаний $T/2$, то расстояние между соседними узлами (или

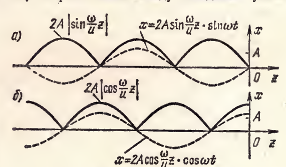


Рис. 11.3. Стоячая волна в струне с закрепленным правым концом (а) и со свободным правым концом (б).

пучностями) стоячей волны равно половине длины бегущей волны $\lambda/2$. График зависимости амплитуды стоячей волны от z показан на рис. 11.3, а. На этом же рисунке пунктиром показано положение струны в некоторый момент времени t.

Из формулы (11.7) видно, что колебания всех точек струны, лежащих между двумя любыми ближайшими узлами, происходят в одинаковой фазе, так как для всех этих точек знак $\sin\frac{\omega z}{u}$ один и тот же. Колебания точек струны, лежащих по разные стороны узла, происходят в противофазе, так как при переходе через узел $\sin\frac{\omega z}{u}$

меняет знак. Фазовые соотношения в стоячей волне хорошо видны из рис. 11.3. Совершенно аналогично рассматривается стоячая волна, возникающая при отражении от свободного конца струны.

Как создать в струне со свободным концом предварительное натяжение, необходимое для распространения волн? Можно, например, прикрепить к правому концу струны легкое кольцо и надеть его на гладкий вертикальный стержень (рис. 11.4). Такое кольцо может скользить с пренебрежимо малым трением по стержню, передавая горизонтальную силу натяжения струне, и не препятствует смещению конца струны в поперечном направлении. Для отраженной от свободного конца струны волны следует взять знак «+» в формуле (11.6). Поэтому уравнение стоячей волны принимает вид

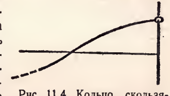


Рис. 11.4. Кольцо, скользящее без трения по стержню, позволяет считать конец струны свободным, хотя струна натянута.

$$x(t,\ z) = 2A\cos\frac{\omega z}{u}\cos\omega t. \qquad (11.9)$$

Картина распределения амплитуд и мгновенное положение струны в такой волне показаны на рис. 11.3, *б*. Узлы и пучности здесь сдвинуты на расстояние $\lambda/4$ по отношению к картине для струны с закрепленным концом.

Находящиеся в узлах стоячей волны частицы струны вообще не движутся. Поэтому через узловые точки не происходит передачи энергии. Стоячая волна, по существу, уже не является волновым движением, хотя и получается в результате интерференции двух бегущих навстречу волн одинаковой амплитуды. То, что стоячая волна уже фактически не волна, а скорее просто колебание, можно увидеть и из энергетических соображений. В бегущей волне кинетическая и потенциальная энергии в каждой точке колеблются в одинаковой фазе. В стоячей волне, как видно, например, из рис. 11.3, колебания кинетической и потенциальной энергий сдвинуты по фазе так же, как и при колебаниях маятника: в тот момент, когда все точки струны одновременно проходят через равновесное положение,

кинетическая энергия струны максимальна, а потенциальная энергия равна нулю.

Попробуем разобраться, что это за колебания. До сих пор мы считали, что в точке $z=0$ находится правый конец струны, свободный или закрепленный, а влево струна тянется до бесконечности. Предположим для определенности, что правый конец струны закреплен и в струне установилась стоячая волна с некоторой длиной λ и амплитудой в пучности, равной $2A$. Она описывается уравнением (11.7).

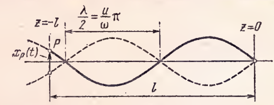


Рис. 11.5. Возбуждение стоячей волны в струне синусоидальным внешним воздействием на ее левый конец.

Теперь представим себе, что мы перерезали струну в некоторой точке $z=-l$ (рис. 11.5), а освободившийся конец приводим в движение внешней силой так, чтобы он совершал точно такие же гармонические колебания, как и в стоячей волне до перерезания струны. Ясно, что движение всех точек струны справа от этой точки при этом никак не изменится. Таким образом, стоячую волну в ограниченной струне длиной l с закрепленным правым концом можно рассматривать как установившееся вынужденное колебание при синусоидальном внешнем воздействии на ее левый конец P. Связь характеристик стоячей волны с законом движения $x_P(t)$ точки P дается выражением, получаемым из формулы (11.7) при подстановке в нее значения $z=-l$:

$$x_P(t)=x(t,\ -l)=-2A\sin\frac{\omega}{u}l\cdot\sin\omega t. \qquad (11.10)$$

Эту формулу можно понимать как условие для выражения амплитуды в пучности стоячей волны $2A$ через амплитуду вынужденных колебаний левого конца струны P. Если записать $x_P(t)$ в виде

$$x_P(t)=-b\sin\omega t, \qquad (11.11)$$

то из сравнения с (11.10) находим

$$2A = \frac{b}{\sin \frac{\omega}{u} l}. \qquad (11.12)$$

Как всегда при вынужденных колебаниях, нас интересует зависимость амплитуды от частоты.

Из формулы (11.12) видно, что амплитуда в пучности стоячей волны будет огромной (в отсутствие затухания — бесконечной) даже при очень малой амплитуде b колебаний левого конца, если $\sin \frac{\omega}{u} l$ обращается в нуль. Но $\sin \frac{\omega}{u} l$ обращается в нуль при такой частоте внешнего воздействия, когда, как видно из рис. 11.5, на левый конец струны P приходится узел стоячей волны. Итак, амплитуда в пучности обращается в бесконечность, если

$$\frac{\omega l}{u} = n\pi, \qquad n = 1,\ 2, \ldots$$

Это значит, что если частота внешней силы совпадает с одной из частот

$$\omega_n = n\frac{\pi u}{l}, \qquad n = 1,\ 2, \ldots \qquad (11.13)$$

то в системе наступает резонанс. Но резонанс в любой системе в отсутствие затухания имеет место при совпадении частоты внешнего воздействия с частотой собственных колебаний. Поэтому набор резонансных частот (11.13) есть в то же время набор частот собственных поперечных колебаний струны длиной l с закрепленными концами. Мы видим, что струна имеет бесконечное число частот собственных колебаний. Это связано с тем, что струна имеет бесконечное число степеней свободы.

Возбудить собственное колебание на частоте ω_n в струне с закрепленными концами можно, например, следующим образом. Придадим струне конфигурацию, соответствующую такой стоячей волне, когда на струне укладывается n полуволн, и отпустим ее без начального толчка (рис. 11.6, *а*). Каждая точка струны будет тогда совершать, как и в стоячей волне, гармоническое колебание с частотой ω_n. Поэтому такое движение струны представляет собой нормальное колебание.

Любое собственное колебание струны можно представить как суперпозицию ее нормальных колебаний. При этом движение каждой точки струны, как и при произвольных колебаниях связанных маятников, уже не будет представлять собой простого гармонического колебания,

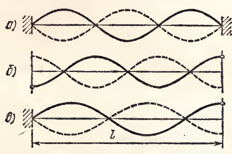


Рис. 11.6. Стоячие волны в струне как ее нормальные моды.

а будет суммой нескольких гармонических колебаний с частотами ω_n.

Легко сообразить, что набор собственных частот струны длиной l со свободными концами совпадает с набором (11.13) для такой же струны с закрепленными концами: в этом случае при нормальном колебании на струне также укладывается целое число полуволн (рис. 11.6, *б*). Если же у струны один конец закреплен, а другой свободен, то при нормальных колебаниях на ней укладывается целое число полуволн и еще четверть волны (рис. 11.6, *в*).

§ 12. Принцип Гюйгенса. Дифракция волн. Эффект Допплера

Наглядное представление о распространении монохроматических волн в упругой среде или на поверхности воды дает картина волновых поверхностей. Напомним, что все точки среды, лежащие на одной волновой поверхности, имеют в данный момент одну и ту же фазу колебания. Другими словами, волновая поверхность — это поверхность постоянной фазы. Уравнение волновой поверхности можно получить, приравнивая фазу в уравнении волны постоянной величине. Например, для плоской волны,

описываемой уравнением

$$x(t,\ z)=A\cos\omega\left(t-\frac{z}{u}\right), \qquad (12.1)$$

уравнение волновой поверхности получаем, приравнивая аргумент косинуса произвольной константе C:

$$\omega\left(t-\frac{z}{u}\right)=C,$$

откуда

$$z=ut+C_1. \qquad (12.2)$$

Видно, что для фиксированного момента времени t уравнение (12.2) — это уравнение плоскости, перпендикулярной оси z. С течением времени эта плоскость перемещается со скоростью u вдоль оси z параллельно самой себе.

Для сферической волны, описываемой уравнением

$$x(t,\ r)=a\frac{r_0}{r}\cos\omega\left(t-\frac{r}{u}\right), \qquad (12.3)$$

поверхность постоянной фазы задается уравнением

$$r=ut+C_2. \qquad (12.4)$$

Волновая поверхность в этом случае — это сфера, центр которой совпадает с центром волны, а радиус r растет с постоянной скоростью u.

Следует различать понятия волновой поверхности и фронта волны. Волновая поверхность введена для монохроматической, строго говоря, бесконечно протяженной волны, при распространении которой все точки среды совершают гармонические колебания. Разумеется, это понятие можно применить и к более общему случаю стационарного волнового процесса, при котором все точки среды совершают периодические (но не обязательно гармонические) колебания по закону $f(t-z/u)$, где $f(t)$ — произвольная периодическая функция времени. Волновые поверхности в этом случае имеют точно такой же вид (12.2), как и в монохроматической волне.

Понятие фронта волны относится к нестационарному волновому процессу распространения возмущения. Пусть вся среда находилась в покое и в некоторый момент времени

включается источник колебания, от которого в среде начинает распространяться возмущение. Фронт волны — это поверхность, которая отделяет точки среды, пришедшие в движение, от тех точек, до которых возмущение еще не дошло. Очевидно, что в однородной изотропной среде фронт волны от плоского источника колебаний представляет собой плоскость, а фронт волны от точечного источника — сферу.

Рис. 12.1. Построение волновой поверхности по принципу Гюйгенса.

При распространении волн в однородной среде нахождение волновых поверхностей не представляет труда. Но при наличии в среде неоднородностей, преград, границ раздела и т. д. нахождение волновых поверхностей усложняется. Простой принцип построения волновых поверхностей был предложен Гюйгенсом. Принцип Гюйгенса позволяет находить волновую поверхность в некоторый момент времени, если известно ее положение в предшествующий момент. Для этого каждую точку волновой поверхности в момент времени t следует рассматривать как источник вторичных волн (рис. 12.1). Волновая поверхность каждой вторичной волны спустя промежуток времени Δt представляет собой в однородной среде сферу радиусом $u\,\Delta t$. Геометрическая огибающая волновых поверхностей вторичных волн и представляет собой искомую волновую поверхность в момент времени $t+\Delta t$. Принцип Гюйгенса можно применять и для нахождения фронта волны в случае нестационарного волнового процесса.

В первоначальной формулировке Гюйгенса этот принцип представлял собой, по существу, лишь удобный рецепт для нахождения волновых поверхностей, ибо он не объяснял, например, то, почему положение волновой поверхности дает именно передняя огибающая вторичных волн и каков смысл задней огибающей поверхности, показанной на рис. 12.1 пунктиром. Обоснование принципа Гюйгенса было дано Френелем на основе учета интерференции вто-

ричных волн. С применением принципа Гюйгенса — Френеля мы встретимся в разделе «Оптика».

Легко видеть, что в простых случаях распространения плоской или сферической волны в однородной среде принцип Гюйгенса приводит к правильным результатам (12.2) и (12.4): плоская волна остается плоской, а сферическая — сферической. Принцип Гюйгенса позволяет найти закон

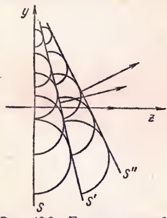


Рис. 12.2. Поворот волновой поверхности в неоднородной среде.

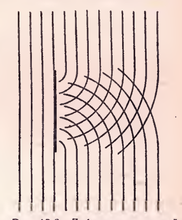


Рис. 12.8. Дифракция плоской волны на поглощающей преграде.

отражения и преломления плоской волны на бесконечной плоской границе раздела двух однородных сред.

С помощью принципа Гюйгенса можно объяснить, почему происходит поворот волновой поверхности при распространении волн в неоднородной среде. Пусть, например, плотность среды ρ возрастает в направлении оси y (рис. 12.2) таким образом, что скорость распространения волн $u(y)$ уменьшается вдоль y по линейному закону. Если в какой-то момент времени t волновая поверхность S представляет собой плоскость $z=0$, то спустя малый промежуток времени, в момент $t+\Delta t$, эта волновая поверхность, как видно из рис. 12.2, поворачивается и занимает новое положение S'. Спустя следующий малый промежуток времени она занимает положение S'', и т. д. Наклоненные друг к другу волновые поверхности должны пересекаться в некоторой точке, где скорость распространения

волн обращается в нуль. Но обращение этой скорости в нуль соответствовало бы бесконечно большой плотности среды, что физически невозможно. Это означает, что на самом деле линейный закон изменения скорости волн может быть осуществлен лишь на ограниченном вдоль оси y участке.

Применение принципа Гюйгенса к распространению волн в среде при наличии преград позволяет качественно объяснить явление дифракции — загибание волн в область геометрической тени. Рассмотрим, например, плоскую волну, падающую на плоскую стенку с прямыми краями (рис. 12.3). Для простоты будем считать, что падающий на стенку участок волны полностью поглощается, так что отраженной волны нет. На рис. 12.3 показаны построенные по принципу Гюйгенса волновые поверхности позади преграды. Видно, что волны действительно загибаются в область тени. Но принцип Гюйгенса ничего не говорит об амплитуде колебаний в волне за преградой. Ее можно найти, рассматривая интерференцию волн, приходящих в область геометрической тени. Распределение амплитуд колебаний позади преграды называется дифракционной картиной. Непосредственно за преградой амплитуда колебаний очень мала. Чем дальше от преграды, тем заметнее становится проникновение колебаний в область геометрической тени.

Полный вид дифракционной картины позади преграды зависит от соотношения между длиной волны λ, размером преграды d и расстоянием L от преграды до точки наблюдения. Если длина волны λ больше размеров препятствия d, то волна его почти не замечает. Если длина волны λ одного порядка с размером преграды d, то дифракция проявляется даже на очень малом расстоянии L и волны за преградой лишь чуть-чуть слабее, чем в свободном волновом поле с обеих сторон. Если, наконец, длина волны много меньше размеров препятствия, то дифракционную картину можно наблюдать только на большом расстоянии от преграды, величина которого зависит от λ и d. Примеры расчета дифракционных картин будут приведены в разделе «Оптика».

Принцип Гюйгенса позволяет найти вид фронта волны для нестационарного волнового процесса, возникающего при движении источника колебаний в неподвижной среде. Здесь возможны два существенно различных случая: ско-

рость источника v меньше скорости распространения волн в среде u и, наоборот, $v>u$. Пусть источник начинает двигаться из точки O по прямой с постоянной скоростью v, непрерывно возбуждая колебания. В первом случае, когда $v<u$, вопрос о форме фронта волны и его положении решается очень просто: фронт волны будет сферическим, а

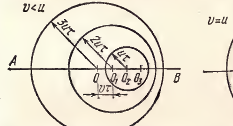


Рис. 12.4. Волновые поверхности при движении источника со скоростью, меньшей скорости волн.

Рис. 12.5. Источник движется с той же скоростью, что и волны.

центр его совпадет с положением источника в начальный момент времени, так как след от всех последующих возмущений окажется внутри этой сферы (рис. 12.4). Действительно, будем рассматривать создаваемые движущимся источником возмущения через равные промежутки времени τ. Точки O_1, O_2 и O_3 дают положение источника в моменты времени τ, 2τ и 3τ. Каждая из этих точек может рассматриваться как центр сферической волны, испущенной источником в тот момент, когда он находился в этой точке. На рис. 12.4 изображены положения фронтов этих волн в момент времени 3τ, когда источник находится в точке O_3. Так как $u>v$, то фронт каждой последующей волны целиком лежит внутри фронта предыдущей.

Если скорость источника равна скорости распространения волн в среде, то, как показано на рис. 12.5, фронты всех волн, испущенных в точках O, O_1 и O_2, соприкасаются в точке O_3, где находится в этот момент источник. Если на фронте каждой волны возникает некоторое уплотнение среды, то непосредственно перед движущимся источником, где фронты всех волн соприкасаются, уплотнение может быть значительным.

Особенно интересен случай, когда скорость источника больше скорости распространения волн в среде: $v>u$. Источник опережает создаваемые им волны. Положение фронтов волн, испущенных в точках O, O_1 и O_2, для того момента времени, когда источник находится в точке O_3,

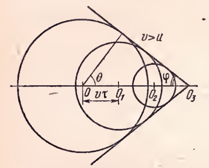


Рис. 12.6. Источник обгоняет создаваемые им волны.

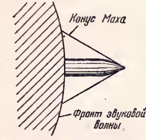


Рис. 12.7. Движущаяся со сверхзвуковой скоростью пуля обгоняет фронт звуковой волны.

показано на рис. 12.6. Огибающая этих фронтов представляет собой поверхность кругового конуса, ось которого совпадает с траекторией источника, вершина в каждый момент совпадает с источником, а угол φ между образующей и осью определяется, как ясно из рис. 12.6, соотношением

$$\sin\varphi=\frac{u}{v}. \qquad (12.5)$$

Такой фронт волны получил название конуса Маха.

С такой формой фронта волны приходится сталкиваться во всех случаях движения тел со сверхзвуковой скоростью — снарядов, ракет, реактивных самолетов. В тех случаях, когда уплотнение среды на фронте волны значительно, фронт волны можно сфотографировать. На рис. 12.7, сделанном по фотографии, показаны конус Маха пули, движущейся со сверхзвуковой скоростью, и фронт звуковой волны, созданной пулей при ее движении в стволе с дозвуковой скоростью. Снимок сделан в тот момент, когда пуля обгоняет фронт звуковой волны.

Аналогом конуса Маха в оптике является черенковское излучение, возникающее при движении заряженных частиц в веществе со скоростью, превышающей скорость света в этой среде.

Из рис. 12.4 видно, что при движении источника монохроматических волн длина излучаемых по разным направлениям волн различна и отличается от длины волны, которую испускал бы неподвижный источник. Если считать

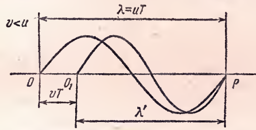


Рис. 12.8. К объяснению эффекта Допплера.

промежуток времени τ равным периоду колебаний $T=2\pi/\omega$, то сферы на рис. 12.4 можно рассматривать как последовательные гребни (или впадины) волн, а расстояние между ними — как длину волны, излучаемой в соответствующем направлении. Видно, что длина волны, излучаемой по направлению движения источника, уменьшается, а в противоположном направлении — увеличивается. Понять, как это происходит, помогает рис. 12.8: источник начинает очередной период излучения волны, находясь в точке O, и, двигаясь в том же направлении, что и волна, заканчивает период, находясь в точке O_1. В результате длина излученной волны λ' оказывается меньше, чем $\lambda=uT$, на величину vT:

$$\lambda' = \lambda - vT = (u-v)\,T = \frac{u-v}{\nu}. \qquad (12.6)$$

Неподвижный приемник, регистрирующий эти волны, будет принимать колебания с частотой ν', отличной от частоты колебаний источника ν:

$$\nu' = \frac{u}{\lambda'} = \frac{u}{u-v}\,\nu. \qquad (12.7)$$

Эта формула справедлива как в случае приближения источника к неподвижному приемнику, так и в случае

удаления. При приближении скорость источника v берется с положительным знаком, при удалении — с отрицательным. Если источник движется с дозвуковой скоростью, то при приближении частота принимаемого звука выше, а при удалении — ниже, чем при неподвижном источнике. Такое изменение высоты звука легко заметить, слушая звук гудка проносящегося мимо поезда или автомобиля. Если скорость приближения источника звука к приемнику стремится к скорости звука, то, согласно (12.6), длина волны λ' стремится к нулю, а частота ν' — к бесконечности.

Если v больше u, то сначала мимо приемника промчится источник и только потом придут созданные им при приближении звуковые волны. Эти волны будут приходить в обратной последовательности по сравнению с тем, как они излучались: волны, излученные раньше, придут позже. В этом смысл отрицательного значения частоты ν', получаемой из формулы (12.7) при $v>u$.

Изменение частоты колебаний, регистрируемых приемником, происходит и в том случае, когда источник волн неподвижен в среде, а движется приемник. Если, например, приемник приближается к источнику со скоростью $v_{пр}$, то его скорость относительно гребней волн равна $u+v_{пр}$. Поэтому регистрируемая им частота колебаний ν' равна

$$\nu' = \frac{u+v_{пр}}{\lambda} = \frac{u+v_{пр}}{u}\nu. \qquad (12.8)$$

Эта формула справедлива и при удалении приемника от неподвижного источника, только скорость $v_{пр}$ нужно взять с отрицательным знаком. Если приемник удаляется от источника со сверхзвуковой скоростью, то он догоняет ранее испущенные волны и регистрирует их в обратной последовательности.

Явление изменения частоты принимаемых волн при движении источника или приемника относительно среды называется эффектом Доплера.

Эффект Доплера имеет место также и для электромагнитных волн. Однако в отличие от звуковых волн, где изменение частоты по-разному зависит от скорости приемника и источника относительно среды, для электромагнитных волн в вакууме эффект Доплера определяется только их относительной скоростью.

§ 13. Волны на воде. Дисперсия и групповая скорость

Во всех рассмотренных нами примерах распространения волн отсутствовала дисперсия: скорость распространения волн не зависела от длины волны. При отсутствии дисперсии возмущение любого вида, которое можно представить как суперпозицию монохроматических волн разной длины, распространяется с такой же скоростью, как и монохроматическая волна, и при этом не меняет своей формы.

Как уже отмечалось, волны на поверхности воды представляют собой более сложный пример волнового движения, чем рассмотренные выше упругие волны в натянутой струне или в однородной упругой среде. Сложность волн на воде проявляется уже в том, что скорость их распространения зависит от длины волны. Непосредственное вычисление скорости таких волн на основе законов динамики затруднительно, поэтому мы попробуем применить для этой цели метод анализа размерностей.

Прежде всего применим этот метод для нахождения скорости волн в уже рассмотренном нами примере распространения продольных волн в бесконечном упругом стержне. От каких свойств стержня она может зависеть? Очевидно, от его упругих свойств, характеризуемых модулем Юнга E, и от инертных свойств, характеризуемых плотностью материала стержня ρ. Как мы уже знаем из динамического рассмотрения, скорость волн не зависит от длины волны λ. Но если бы мы сейчас впервые приступали к решению такой задачи, то должны были бы допустить возможность такой зависимости. Можно было бы допустить и зависимость скорости волн от их амплитуды, однако для волн малой амплитуды, когда $A \ll \lambda$, такой зависимости нет. Только такие волны мы и будем рассматривать.

Итак, параметры, характеризующие распространение волн в стержне,— это E, ρ, λ. Напомним размерности этих величин:

$$[E] = ML^{-1}T^{-2}, \quad [\rho] = ML^{-3}, \quad [\lambda] = L. \qquad (13.1)$$

Легко убедиться, что из этих параметров безразмерную комбинацию составить невозможно. Действительно, размерность модуля Юнга E содержит время, которого нет в размерностях остальных величин,— значит, модуль Юнга

не может входить в безразмерную комбинацию. Но из двух оставшихся величин ρ и λ также нельзя составить безразмерную комбинацию, так как в размерность ρ входит масса, а в размерность λ — нет.

Составим из E, ρ и λ величину, имеющую размерность скорости:

$$u = CE^x \rho^y \lambda^z. \tag{13.2}$$

Этому выражению соответствует следующее равенство размерностей:

$$LT^{-1} = (ML^{-1}T^{-2})^x (ML^{-3})^y L^z. \tag{13.3}$$

Система уравнений для нахождения x, y и z имеет вид

$$\begin{array}{ll} L & -x-3y+z=1, \\ T & \qquad -2x=-1, \\ M & \qquad x+y=0, \end{array}$$

откуда $x=1/2$, $y=-1/2$, $z=0$. Итак, согласно (13.2) выражение для скорости волн u имеет вид

$$u = C\sqrt{\frac{E}{\rho}}. \tag{13.4}$$

Мы видим, что, как и при динамическом рассмотрении, скорость волн не зависит от длины волны. Метод размерностей, разумеется, не дает возможности определить значение численного коэффициента C, но дает правильную зависимость скорости волн от свойств среды.

Перейдем теперь к волнам на воде. Возможность распространения волн на поверхности воды обусловлена действием поля тяжести и сил поверхностного натяжения. Роль этих сил различна для волн разной длины: для достаточно коротких волн, когда кривизна поверхности жидкости велика, преобладающими являются силы поверхностного натяжения, а в случае длинных волн этими силами, наоборот, можно пренебречь. В первом случае волны на воде называются капиллярными и представляют собой мелкую рябь. Во втором случае волны называются гравитационными. Эти случаи следует рассматривать отдельно.

Начнем с капиллярных волн. Параметрами, от которых может зависеть скорость таких волн, являются коэффициент

поверхностного натяжения σ, плотность воды ρ и длина волны λ. Размерность σ есть MT^{-2}. Опять легко видеть, что безразмерную комбинацию из σ, ρ и λ составить нельзя. Составляем из σ, ρ и λ комбинацию, имеющую размерность скорости. Обозначив скорость капиллярных волн через u_σ, можем написать

$$u_\sigma = C\sigma^x\rho^y\lambda^z. \tag{13.5}$$

Соответствующее (13.5) равенство размерностей имеет вид

$$LT^{-1} = (MT^{-2})^x (ML^{-3})^y L^z. \tag{13.6}$$

Приравнивая показатели степеней при L, T и M, находим: $x=1/2$, $y=z=-1/2$. Поэтому

$$u_\sigma = C\sqrt{\frac{\sigma}{\rho\lambda}}. \tag{13.7}$$

Точная динамическая теория дает для C значение, равное $\sqrt{2\pi}$.

Скорость распространения капиллярных волн оказалась зависящей не только от свойств среды, характеризуемых поверхностным натяжением σ и плотностью ρ, но и от длины волны λ. Значит, для капиллярных волн имеет место дисперсия.

Рассмотрим теперь гравитационные волны. Так как обычно это волны большого масштаба, то естественно предположить, что их скорость может зависеть и от глубины водоема h. Поэтому параметры, определяющие скорость распространения гравитационных волн на поверхности воды u_g, суть g, ρ, λ, h. Из этих величин, как легко убедиться, можно составить только одну независимую безразмерную комбинацию λ/h. В выражение для скорости волн не может входить величина ρ, ибо ее размерность содержит массу, которой нет в размерностях остальных величин. Это и понятно, так как и вызывающая колебания воды сила тяжести, и инертные свойства воды пропорциональны одной и той же величине — ее плотности. Итак, выражение для скорости u_g имеет вид

$$u_g = \sqrt{g\lambda}\, f\left(\frac{\lambda}{h}\right), \tag{13.8}$$

где f — произвольная функция безразмерного параметра λ/h, вид которой не может быть определен из соображений

размерности. Как обычно, вид функции f несложно установить в предельных случаях с помощью дополнительных физических соображений. Ясно, что на очень глубокой воде, когда $\lambda \ll h$, скорость волны не может зависеть от глубины водоема. Поэтому $f(\lambda/h)$ в этом случае стремится

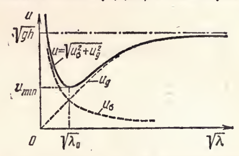


Рис. 13.1. Скорость волн на поверхности воды.

к некоторой постоянной величине C и выражение (13.8) принимает вид

$$u_g = C\sqrt{g\lambda}, \qquad \lambda \ll h. \tag{13.9}$$

Динамическая теория дает для C значение, равное $(2\pi)^{-1/2}$. Видно, что для гравитационных волн на глубокой воде, как и для капиллярных волн, имеет место дисперсия, хотя зависимость скорости от длины волны в этом случае иная. В другом предельном случае — волны на очень мелкой воде, когда $h \ll \lambda$,— скорость распространения волны не должна зависеть от длины волны. Поэтому из (13.8) следует, что $f(\lambda/h)$ для волн на мелкой воде должна содержать $\sqrt{\lambda}$ в знаменателе: $f(\lambda/h) = C\sqrt{h/\lambda}$. Выражение для скорости волн u_g принимает вид

$$u_g = C\sqrt{gh}, \qquad h \ll \lambda. \tag{13.10}$$

Динамическая теория дает в этом случае $C=1$. Скорость распространения предельно длинных гравитационных волн зависит от глубины водоема h: чем глубже вода, тем больше скорость распространения.

Построим зависимость скорости распространения от длины волны для волн разного вида на поверхности воды. Удобно по оси абсцисс откладывать не λ, а $\sqrt{\lambda}$. Скорость капиллярных волн изображается гиперболой на рис. 13.1.

Скорость гравитационных волн изображается прямой, которая по мере роста длины волны загибается и асимптотически приближается к прямой $u=\sqrt{gh}$. Скорость капиллярных волн падает с увеличением длины волны, а скорость гравитационных растет, приближаясь к максимальному значению, равному $\sqrt{gh}$. Кривые u_σ и u_g пересекаются при некотором значении $\lambda=\lambda_0$. При $\lambda \ll \lambda_0$ волны на воде являются чисто капиллярными, при $\lambda \gg \lambda_0$ — чисто гравитационными. Если же длина волны λ близка к λ_0, то распространение таких волн определяется в равной мере и поверхностным натяжением, и силой тяжести. Динамическое рассмотрение показывает, что скорость распространения таких капиллярно-гравитационных волн u дается выражением

$$u=\sqrt{u_\sigma^2+u_g^2}. \tag{13.11}$$

Зависимость u от $\sqrt{\lambda}$ показана сплошной кривой на рис. 13.1. Минимум этой кривой приходится на значение $\lambda=\lambda_0$. Определим λ_0, приравнивая скорости u_σ и u_g. Используя значения $\sigma=72$ дин/см, $g=980$ см/с^2 и $\rho=1$ г/см^3, получим $\lambda_0=1{,}73$ см. Минимальное значение скорости u_{min} определяется с помощью (13.11) и оказывается равным 23,2 см/с. На поверхности воды не может существовать волн, распространяющихся со скоростью, меньшей 23,2 см/с!

Рис. 13.2. Профиль гравитационных волн на воде.

Капиллярные волны на поверхности воды — это рябь с длиной волн, меньшей 1 см. Те волны, которые мы обычно видим на поверхности воды, имеют гораздо бóльшую длину волны и являются гравитационными. Профиль таких волн лишь в предельном случае очень малых амплитуд напоминает синусоиду. Волны на воде имеют узкие острые гребни и широкие плоские впадины (рис. 13.2). Такая форма волны обусловлена тем, что волна на поверхности воды не является чисто поперечной: отдельные частицы жидкости при прохождении волны движутся по замкнутым траекториям, близким к круговым.

Скорость гравитационных волн на глубокой воде растет с увеличением длины волны. Этот рост, как видно из

рис. 13.1, замедляется, как только длина волны становится сравнимой с глубиной водоема. Поэтому максимальная скорость распространения гравитационных волн на воде определяется глубиной водоема. Наибольшая глубина встречается в открытом океане. Там и следует ожидать самых быстрых волн. Оценим их максимальную скорость. Примем глубину океана равной 5 км. Наибольшую скорость будут иметь волны, длина которых значительно больше 5 км. Для них даже такая вода будет «мелкой», и с помощью формулы (13.10) находим $u \approx 200$ м/с, т. е. примерно 700 км/ч,— волна бежит со скоростью самолета. Столь длинные волны возникают при подводных землетрясениях и называются цунами.

Как мы видели, скорость распространения волн на воде оказалась зависящей от длины волны, т. е. имеет место дисперсия. Предположим, что на поверхности воды распространяется не отдельная монохроматическая волна бесконечной протяженности, а группа волн, представляющая собой цуг ограниченной длины. С какой скоростью будет распространяться центр такого цуга? Представление о движении цуга волн можно получить, рассматривая волну, образующуюся при сложении двух монохроматических волн с близкими длинами λ и $\lambda+\Delta\lambda$. В отсутствие дисперсии эти волны распространялись бы с одинаковой скоростью u. При наличии дисперсии они распространяются с несколько различающимися скоростями u и $u+\Delta u$.

Как выглядит моментальная «фотография» результирующей волны? Фотографии каждой из складываемых волн представляли бы собой застывшие синусоиды с разной длиной волны (рис. 9.5). В том месте, где горб одной из этих волн совпадает с горбом другой, результирующая волна имеет горб удвоенной высоты. Там, где горб одной волны совпадает с впадиной другой, в результирующей волне смещение равно нулю. Как видно из рис. 9.5, «фотография» результирующей волны представляет собой последовательность отдельных групп волн. Как вся эта картина меняется со временем?

Если скорости складываемых волн одинаковы, то результирующая волна распространяется с той же скоростью, не изменяя своей формы. Если же скорости складываемых волн различаются, то взаимное расположение их горбов и впадин меняется с течением времени. Мгновенная «фо-

тография» результирующей волны будет, разумеется, иметь такой же вид, как и раньше, но положение центров отдельных групп волн с течением времени будет изменяться относительно горбов и впадин складываемых волн. Поэтому центры отдельных групп волн движутся с иной скоростью, нежели складываемые синусоидальные волны. Скорость движения центров этих групп называют групповой скоростью. Найдем эту скорость.

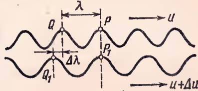


Рис. 13.3. К выводу формулы для групповой скорости волн.

Будем для определенности считать, что скорость монохроматических волн растет с увеличением длины волны. Тогда нижняя волна на рис. 13.3, имеющая длину $\lambda+\Delta\lambda$, обгоняет верхнюю волну с длиной λ. Пусть в какой-то момент времени совпадают горбы P и P_1, т. е. центр группы волн приходится на точку P. Через некоторое время τ горб P_1 обгонит P, но зато совпадут горбы Q и Q_1. Это значит, что центр группы волн за это время сместился назад на одну длину волны λ и совпадает с точкой Q. Поэтому скорость перемещения центра группы волн в пространстве $u_г$ меньше скорости верхней волны на величину λ/τ:

$$u_г = u - \frac{\lambda}{\tau}. \qquad (13.12)$$

Время τ, в течение которого горб Q_1 догоняет Q, как легко видеть из рис. 13.3, равно $\Delta\lambda/\Delta u$. Поэтому выражение для групповой скорости (13.12) в пределе при $\Delta\lambda\to 0$ принимает вид

$$u_г = u - \lambda\frac{du}{d\lambda}. \qquad (13.13)$$

Вычислим групповую скорость для волны на поверхности воды. Для гравитационных волн на глубокой воде с помощью формулы (13.9) получаем

$$\frac{du_g}{d\lambda} = \frac{d}{d\lambda}(C\sqrt{g\lambda}) = \frac{1}{2}C\sqrt{\frac{g}{\lambda}} = \frac{1}{2}\frac{u_g}{\lambda}. \qquad (13.14)$$

Подставляя это значение в (13.13), находим

$$u_г = u_g - \frac{1}{2} u_g = \frac{1}{2} u_g. \qquad (13.15)$$

Скорость распространения центра группы гравитационных волн на глубокой воде оказывается вдвое меньше скорости монохроматических волн. На мелкой воде групповая скорость гравитационных волн совпадает со скоростью монохроматических волн, так как для них отсутствует дисперсия ($du/d\lambda = 0$).

Для нахождения групповой скорости капиллярных волн вычислим $du/d\lambda$ с помощью формулы (13.7):

$$\frac{du_\sigma}{d\lambda} = \frac{d}{d\lambda}\left(C\sqrt{\frac{\sigma}{\rho\lambda}}\right) = -\frac{C}{2\lambda}\sqrt{\frac{\sigma}{\rho\lambda}} = -\frac{u_\sigma}{2\lambda}. \qquad (13.16)$$

Подставляя в (13.13), получаем

$$u_г = u_\sigma + \frac{u_\sigma}{2} = \frac{3}{2} u_\sigma. \qquad (13.17)$$

Центр группы капиллярных волн бежит в полтора раза быстрее, чем отдельная монохроматическая волна.

Каждая группа волн, как видно из рис. 9.5, состоит из горбов и впадин, которые движутся с такой же скоростью, что и монохроматическая волна. Мы видим, что при наличии дисперсии группа как целое движется с иной скоростью, чем входящие в ее состав горбы и впадины. Как можно это себе представить? Так получается, потому что в процессе распространения группа «живет»: на одном конце группы возникают новые горбы, а на другом — горбы угасают.

§ 14. Электромагнитные волны

Основные закономерности волновых процессов носят универсальный характер и в равной мере справедливы для волн различной физической природы: механических волн в упругой среде, волн на поверхности воды, в натянутой струне и т. п. Не являются исключением и электромагнитные волны, представляющие собой процесс распространения колебаний электромагнитного поля.

Но в отличие от всех других видов волн, распространение которых всегда происходит в какой-то среде, элект-

ромагнитные волны могут распространяться в пустоте: никакой материальной среды для распространения электрического и магнитного полей не требуется. Но, разумеется, электромагнитные волны могут существовать не только в вакууме, но и в веществе.

Существование электромагнитных волн было теоретически предсказано Максвеллом в результате анализа предложенной им системы уравнений, описывающих электромагнитное поле. Максвелл показал, что электромагнитное поле в вакууме может существовать и в отсутствие источников — зарядов и токов. Поле без источников имеет вид волн, распространяющихся с конечной скоростью $c=3\cdot 10^{10}$ см/с, в которых векторы электрического и магнитного полей в каждый момент времени в каждой точке пространства перпендикулярны друг другу и перпендикулярны направлению распространения волн.

Экспериментально электромагнитные волны были открыты и изучены Герцем только спустя 10 лет после смерти Максвелла.

Совпадение скорости электромагнитных волн с измеренной задолго до их открытия скоростью света послужило отправным пунктом для отождествления света с электромагнитными волнами и создания электромагнитной теории света.

Электромагнитная волна существует без источников полей в том смысле, что после ее излучения электромагнитное поле волны не связано с источником. Этим электромагнитная волна отличается от статических электрического и магнитного полей, которые не существуют в отрыве от источника.

Излучение электромагнитных волн происходит при ускоренном движении электрических зарядов. Понять, каким образом поперечное электрическое поле волны возникает из радиального кулоновского поля точечного заряда, можно с помощью следующего простого рассуждения, предложенного Дж. Томсоном.

Рассмотрим электрическое поле, создаваемое точечным зарядом q. Если заряд покоится, то его электростатическое поле изображается радиальными силовыми линиями, выходящими из заряда (рис. 14.1). Пусть в момент времени $t=0$ заряд под действием какой-то внешней силы начинает двигаться с ускорением a, а спустя некоторое время τ

действие этой силы прекращается, так что дальше заряд движется равномерно со скоростью $v=a\tau$. График скорости движения заряда показан на рис. 14.2.

Представим себе картину силовых линий электрического поля, создаваемого этим зарядом, спустя большой

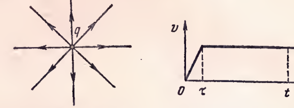


Рис. 14.1. Поле неподвижного точечного заряда.

Рис. 14.2. График скорости заряда.

промежуток времени t: $t \gg \tau$. Поскольку электрическое поле распространяется со скоростью света c, то до точек, лежащих за пределами сферы радиуса ct, изменение электрического поля, вызванное движением заряда, дойти не могло: за пределами этой сферы поле такое же, каким оно было при неподвижном заряде (рис. 14.3). Напряженность этого поля равна

Рис. 14.3. Силовые линии электрического поля заряда, движущегося согласно графику на рис. 14.2.

$$E(r)=\frac{1}{4\pi\varepsilon_0}\frac{q}{r^2}. \quad (14.1)$$

Все изменение электрического поля, вызванное ускоренным движением заряда в течение времени τ, в момент времени t находится внутри тонкого шарового слоя толщиной $c\tau$, наружный радиус которого равен ct, а внутренний $c(t-\tau)$. Это показано на рис. 14.3. Внутри сферы радиуса $c(t-\tau)$ электрическое поле — это поле равномерно движущегося заряда. Если скорость заряда v много

меньше скорости света c, то это поле в момент времени t совпадает с полем неподвижного точечного заряда q, находящегося на расстоянии vt от начала (рис. 14.3): поле медленно движущегося с постоянной скоростью заряда перемещается вместе с ним, а пройденное зарядом за время t расстояние, как видно из рис. 14.2, можно считать равным vt, если $t \gg \tau$.

Картину электрического поля внутри шарового слоя легко найти, учитывая непрерывность силовых линий. Для этого нужно соединить соответствующие радиальные силовые линии (рис. 14.3). Вызванный ускоренным движением заряда излом силовых линий убегает от заряда со скоростью c. Изломы на силовых линиях между сферами $r=ct$ и $r=c(t-\tau)$ — это и есть интересующее нас поле излучения, распространяющееся со скоростью c.

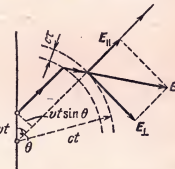


Рис. 14.4. К выводу формулы для напряженности поля излучения ускоренно движущегося заряда.

Чтобы найти поле излучения, рассмотрим одну из силовых линий, составляющую некоторый угол θ с направлением движения заряда (рис. 14.4). Разложим вектор напряженности электрического поля в изломе $\boldsymbol{E}$ на две составляющие: радиальную $\boldsymbol{E}_{\parallel}$ и поперечную $\boldsymbol{E}_{\perp}$. Радиальная составляющая $\boldsymbol{E}_{\parallel}$ — это напряженность электростатического поля, создаваемого зарядом q на расстоянии $r=ct$ от него:

$$E_{\parallel}=\frac{1}{4\pi\varepsilon_0}\frac{q}{(ct)^2}. \qquad (14.2)$$

Поперечная составляющая $\boldsymbol{E}_{\perp}$ — это напряженность электрического поля в волне, излученной зарядом при ускоренном движении. Так как эта волна бежит по радиусу, то вектор $\boldsymbol{E}_{\perp}$ перпендикулярен направлению распространения волны. Из рис. 14.4 видно, что

$$\frac{E_{\perp}}{E_{\parallel}}=\frac{vt\sin\theta}{c\tau}. \qquad (14.3)$$

Подставляя сюда $E_{\parallel}$ из (14.2), находим

$$E_{\perp}=\frac{1}{4\pi\varepsilon_0}\frac{qv}{c^3t\tau}\sin\theta.$$

Учитывая, что $ct=r$, а отношение v/τ есть ускорение a, с которым двигался заряд в течение промежутка времени от 0 до τ, перепишем это выражение в виде

$$E_{\perp}=\frac{1}{4\pi\varepsilon_0}\frac{qa}{c^2r}\sin\theta. \qquad (14.4)$$

Прежде всего обратим внимание на то, что напряженность электрического поля волны $E_{\perp}$ убывает обратно пропорционально расстоянию r от центра, в отличие от напряженности электростатического поля $E_{\parallel}$, которая пропорциональна $1/r^2$. Такой зависимости от расстояния и следовало ожидать, если принять во внимание закон сохранения энергии. Так как при распространении волны в пустоте поглощения энергии не происходит, то количество энергии, прошедшее через сферу любого радиуса, одинаково. Поскольку площадь поверхности сферы пропорциональна квадрату ее радиуса, то поток энергии через единицу ее поверхности должен быть обратно пропорционален квадрату радиуса. Учитывая, что плотность энергии электрического поля волны равна $\frac{1}{2}\varepsilon_0E_{\perp}^2$, приходим к выводу, что $E_{\perp}\sim 1/r$.

Далее отметим, что напряженность поля волны $E_{\perp}$ в формуле (14.4) в момент времени t зависит от ускорения заряда a в момент времени $t=0$: волна, излученная в момент $t=0$, достигает точки, находящейся на расстоянии r, спустя время, равное r/c.

Предположим теперь, что заряд q все время движется вдоль прямой с некоторым переменным ускорением $a(t)$ вблизи начала координат, например совершает гармонические колебания. Тогда он будет излучать электромагнитные волны непрерывно. Напряженность электрического поля волны в точке, находящейся на расстоянии r от начала координат, по-прежнему определяется формулой (14.4), причем поле $E_{\perp}$ в момент времени t зависит от ускорения заряда a в более ранний момент $t-r/c$:

$$E_{\perp}(t)=\frac{1}{4\pi\varepsilon_0}\frac{qa\left(t-\frac{r}{c}\right)}{c^2r}\sin\theta, \qquad (14.5)$$

Пусть движение заряда представляет собой гармоническое колебание вблизи начала координат с некоторой амплитудой A и частотой ω:

$$x(t)=A\cos\omega t. \qquad (14.6)$$

Ускорение заряда $a(t)$ при таком движении дается выражением

$$a(t)=\ddot{x}(t)=-\omega^2 A\cos\omega t. \qquad (14.7)$$

Подставляя ускорение заряда $a(t)$ в формулу (14.5), получаем

$$E_\perp(t)=-\frac{1}{4\pi\varepsilon_0}\frac{q\omega^2 A\cos\omega\left(t-\frac{r}{c}\right)}{c^2 r}\sin\theta. \qquad (14.8)$$

Изменение электрического поля в любой точке при прохождении такой волны представляет собой гармоническое колебание с частотой ω, т. е. осциллирующий заряд излучает монохроматическую волну. Разумеется, формула (14.8) справедлива на расстояниях r, больших по сравнению с амплитудой колебаний заряда A.

Плотность энергии электрического поля $w_э$ монохроматической волны, излучаемой зарядом, можно найти с помощью формулы (14.8):

$$w_э=\frac{1}{2}\varepsilon_0 E_\perp^2=\frac{1}{32\pi^2\varepsilon_0}\frac{q^2\omega^4 A^2\cos^2\omega\left(t-\frac{r}{c}\right)}{c^4 r^2}\sin^2\theta. \qquad (14.9)$$

Плотность энергии пропорциональна квадрату амплитуды колебаний заряда и четвертой степени частоты.

Любое колебание связано с периодическими переходами энергии из одного вида в другой и обратно. Например, колебания механического осциллятора сопровождаются взаимными превращениями кинетической энергии и потенциальной энергии упругой деформации. При изучении электромагнитных колебаний в контуре мы видели, что аналогом потенциальной энергии механического осциллятора является энергия электрического поля в конденсаторе, а аналогом кинетической энергии — энергия магнитного поля катушки. Эта аналогия справедлива не только для локализованных колебаний, но и для волновых процессов. В монохроматической волне, бегущей в упругой

среде, плотности кинетической и потенциальной энергий, как мы видели, в каждой точке совершают гармоническое колебание с удвоенной частотой, причем так, что их значения совпадают в любой момент времени. Так же и в бегущей монохроматической электромагнитной волне: плотности энергии электрического и магнитного полей, совершая гармоническое колебание с частотой 2ω, равны друг другу в каждой точке в любой момент времени.

Плотность энергии магнитного поля $w_м$ выражается через индукцию B, как было показано в разделе «Электричество и магнетизм», следующим образом:

$$w_м = \frac{1}{2}\frac{B^2}{\mu_0}$$

(формула (9.12) упомянутого раздела). Приравнивая плотности энергии электрического и магнитного полей в бегущей электромагнитной волне, убеждаемся, что индукция магнитного поля в такой волне зависит от координат и времени точно так же, как напряженность электрического поля. Другими словами, в бегущей волне индукция магнитного поля $B_\perp$ и напряженность электрического поля $E_\perp$ пропорциональны друг другу в любой точке в любой момент времени:

$$B_\perp = \sqrt{\varepsilon_0\mu_0}E_\perp = \frac{1}{c}E_\perp.$$

Полная плотность энергии электромагнитного поля в бегущей волне w вдвое больше плотности энергии электрического поля (14.9). Плотность потока энергии j, переносимой волной, равна произведению плотности энергии w на скорость распространения волны c: $\vec{j} = c w$. С помощью формулы (14.9) можно увидеть, что поток энергии через любую поверхность осциллирует с частотой 2ω. Для нахождения среднего значения плотности потока энергии $\langle j \rangle$ необходимо усреднить по времени выражение (14.9). Так как среднее значение $\cos^2\omega\left(t-\frac{r}{c}\right)$ равно 1/2, то для $\langle j \rangle$ получаем

$$\langle j \rangle = \frac{1}{32\pi^2\varepsilon_0}\frac{q^2\omega^4A^2}{c^3r^2}\sin^2\theta. \qquad (14.10)$$

Плотность потока энергии в волне зависит от направления: в том направлении, по которому происходят колебания

заряда, энергия вовсе не излучается ($\sin\theta=0$). Наибольшее количество энергии излучается в плоскости, перпендикулярной этому направлению ($\sin\theta=1$). Угловое распределение излучаемой осциллирующим зарядом энергии показано на рис. 14.5. Заряд совершает колебания вдоль оси z. Из начала координат проводятся отрезки, длина которых пропорциональна излучаемой в данном направлении энергии, т. е. $\sin^2\theta$. На диаграмме показана линия, соединяющая концы этих отрезков.

Рис. 14.5. Угловое распределение энергии, излучаемой осциллирующим зарядом.

Распределение энергии по направлениям в пространстве характеризуется поверхностью, которая получается вращением диаграммы вокруг оси z.

Электромагнитная волна в вакууме является поперечной: вектор напряженности электрического поля волны, как это видно из приведенных выше рассуждений, перпендикулярен направлению распространения волны. Проведем через точку наблюдения P на рис. 14.6 сферу с центром в начале координат, около которого вдоль оси z совершает колебания излучающий заряд. Проведем на ней параллели и меридианы. Тогда вектор $\boldsymbol{E}$ поля волны направлен по касательной к меридиану, а вектор $\boldsymbol{B}$ перпендикулярен вектору $\boldsymbol{E}$ и направлен по касательной к параллели. Чтобы убедиться в этом, рассмотрим подробнее взаимосвязь электрического и магнитного полей в бегущей волне. Эти поля после излучения волны уже не связаны с источником. При изменении электрического поля волны возникает магнитное поле, силовые линии которого, как мы видели при изучении тока смещения, перпендикулярны силовым линиям электрического поля. Это переменное магнитное поле, изменяясь, в свою очередь приводит к появлению

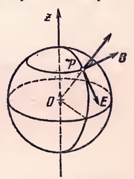


Рис. 14.6. Направления векторов $\boldsymbol{E}$ и $\boldsymbol{B}$ в волне, излучаемой осциллирующим вдоль оси z зарядом.

вихревого электрического поля, которое перпендикулярно породившему его магнитному полю. Таким образом, при распространении волны электрическое и магнитное поля поддерживают друг друга, оставаясь все время взаимно перпендикулярными. Так как в бегущей волне изменение электрического и магнитного полей происходит в фазе друг с другом, то мгновенный «портрет» волны (векторы

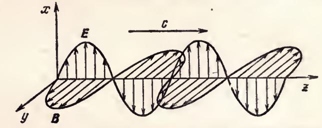


Рис. 14.7. Электрическое и магнитное поля в бегущей линейно поляризованной волне.

E и ***B*** в разных точках линии вдоль направления распространения) имеет вид, показанный на рис. 14.7. Такая волна называется линейно поляризованной. Совершающий гармоническое колебание заряд излучает по всем направлениям линейно поляризованные волны. В бегущей по любому направлению линейно поляризованной волне вектор ***E*** все время находится в одной плоскости.

Возможны и другие виды поляризации поперечных электромагнитных волн. Если, например, вектор ***E*** в некоторой точке при прохождении волны равномерно вращается вокруг направления распространения, оставаясь неизменным по величине, то волна называется циркулярно поляризованной или поляризованной по кругу. Мгновенный «портрет» электрического поля такой волны показан на рис. 14.8. Волну круговой поляризации можно получить при сложении двух распространяющихся в одном направлении линейно поляризованных волн одинаковой частоты и амплитуды, векторы электрического поля в которых взаимно перпендикулярны. В каждой волне вектор электрического поля в каждой точке совершает гармоническое колебание. Чтобы при сложении таких взаимно перпендикулярных колебаний получилось вращение результирующего вектора, необходим сдвиг фаз на $\pi/2$. Другими сло-

вами, складываемые линейно поляризованные волны должны быть сдвинуты на четверть длины волны друг относительно друга.

Наряду с энергией электромагнитная волна обладает и импульсом. Если волна поглощается, то ее импульс

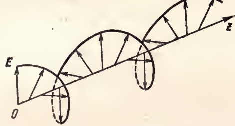


Рис. 14.8. Электрическое поле в бегущей циркулярно поляризованной волне.

передается тому объекту, который ее поглощает. Отсюда следует, что при поглощении электромагнитная волна оказывает давление на преграду. Объяснить происхождение давления волны и найти величину этого давления можно следующим образом.

Падающая волна взаимодействует с электрическими зарядами, входящими в состав любого тела. Сила, с которой электрическое и магнитное поля волны действуют на заряд q, равна

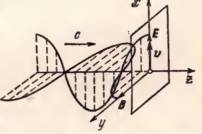


Рис. 14.9. Падение линейно поляризованной волны на поглощающую преграду.

$$F = qE + qv \times B, \quad (14.11)$$

где v — скорость движения этого заряда. В бегущей электромагнитной волне в любой момент времени $B = E/c$; поэтому при $v \ll c$ второе слагаемое в (14.11) — сила Лоренца — всегда много меньше первого. Но именно с силой Лоренца связано давление, оказываемое волной, в то время как первое слагаемое, qE, определяет энергию, поглощаемую преградой. На рис. 14.9 показано падение линейно поля-

ризованной волны на преграду: ось z выбрана по направлению распространения волны, ось x — вдоль направления колебаний вектора $\boldsymbol{E}$. Будем считать, что движение заряда в поглощающей волну преграде вызвано электрическим полем волны и поэтому векторы $\boldsymbol{E}$ и $\boldsymbol{v}$ направлены по одной прямой. Тогда поглощаемая зарядом мощность P равна

$$P = \boldsymbol{F}\cdot\boldsymbol{v} = q\boldsymbol{E}\cdot\boldsymbol{v} = qE_x v_x. \qquad (14.12)$$

Действующая на заряд сила Лоренца перпендикулярна векторам $\boldsymbol{v}$ и $\boldsymbol{B}$ и направлена по оси z:

$$F_z = qv_x B_y. \qquad (14.13)$$

Так как $B_y(t) = E_x(t)/c$, то

$$F_z = \frac{qE_x v_x}{c}.$$

Сравнивая это выражение с выражением для поглощаемой мощности (14.12), находим

$$F_z = \frac{1}{c}P. \qquad (14.14)$$

Так как взаимодействие волны с зарядом q в среднем приводит к поглощению энергии волны, то среднее за период значение поглощаемой мощности положительно. Поэтому и действующая на заряд средняя сила направлена в ту же сторону, что и волна.

Будем считать, что вся энергия падающей волны поглощается преградой. Так как на единицу площади поверхности преграды в единицу времени волна приносит энергию $j = cw$, то оказываемое волной при нормальном падении давление p равно плотности энергии волны w: $p = w$. Сила давления поглощаемой электромагнитной волны сообщает преграде в единицу времени импульс, равный, согласно формуле (14.14), поглощенной энергии, деленной на скорость света c. А это означает, что поглощенная электромагнитная волна обладала импульсом, величина которого равна энергии, деленной на скорость света.

Впервые давление электромагнитных волн экспериментально было обнаружено П. Н. Лебедевым в 1900 году в исключительно тонких опытах.

5. ОПТИКА. ТЕОРИЯ ОТНОСИТЕЛЬНОСТИ

ВОЛНОВАЯ ОПТИКА

§ 1. Свет как электромагнитные волны. Интерференция

Оптика — это учение о физических явлениях, связанных с распространением коротких электромагнитных волн, длина которых составляет приблизительно 10^{-6}—10^{-7} м. Значение именно этой области спектра электромагнитных волн связано с тем, что внутри нее в узком интервале длин волн от 0,40 до 0,76 мкм лежит участок видимого света, непосредственно воспринимаемого человеческим глазом. С точки зрения физики происходящих процессов выделение столь узкой области видимого света не имеет особого смысла, поэтому в понятие «оптический диапазон» включают обычно еще и инфракрасное и ультрафиолетовое излучение. Но и для них принятые границы спектра достаточно условны. По существу, эти границы определяются используемыми способами получения и регистрации электромагнитных волн.

Излучение электромагнитных волн происходит при ускоренном движении электрических зарядов. Электромагнитные волны радиодиапазона излучаются антеннами радиопередатчиков при вынужденных колебаниях электронов в антеннах. Все электроны в антенне совершают колебания в одинаковой фазе. Поскольку эти колебания могут поддерживаться очень долго и с высоким постоянством частоты, то излучаемые при этом радиоволны с огромной степенью точности можно считать монохроматическими.

В оптике все иначе. Любой источник света — это скопление множества возбужденных или непрерывно возбуждаемых атомов. Генератор световой волны — это каждый

отдельный атом вещества. Возбужденный атом излучает цуг почти монохроматических волн конечной протяженности. Характерной особенностью каждого элементарного источника является его самостоятельность, независимость от других атомов. Поэтому даже в том случае, когда отдельные цуги можно характеризовать одной и той же длиной волны λ, соотношения фаз между цугами волн, излученных разными атомами, имеют совершенно случайный характер и непрерывно меняются. Только в лазере, где используется вынужденное излучение, удается заставить все возбужденные атомы излучать электромагнитные волны согласованно, подобно тому как это происходит в антенне радиопередатчика. В результате образуется световая волна, близкая по своим свойствам к идеальной монохроматической,— когерентная электромагнитная волна. Излучение обычных источников света, таких, как раскаленные твердые или жидкие тела, возбужденные электрическим разрядом газы и т. д., представляет собой наложение огромного числа не согласованных между собой цугов волн, т. е. фактически «световой шум» — беспорядочные, некогерентные колебания электромагнитного поля.

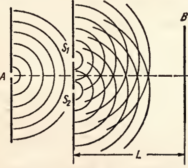


Рис. 1.1. Схема опыта Юнга.

Наблюдать интерференцию света от таких некогерентных источников можно, только используя специальные приемы — разделяя исходный пучок на два. Хотя в каждом из этих пучков, как и в исходном, фазовые соотношения между различными цугами непрерывно хаотически меняются, эти изменения будут одинаковыми для обоих пучков. Если эти пучки снова свести вместе, то можно наблюдать устойчивую интерференционную картину при условии, что разность хода между пучками не превышает длины отдельного цуга. Если же разность хода окажется больше длины цуга, то устойчивой интерференционной картины не будет, так как в этом случае будет происходить наложение цугов, излученных разными атомами.

Явление интерференции света впервые было объяснено на основе волновых представлений Юнгом в 1802 году. В произведенном им опыте малое отверстие A в непрозрачном экране (рис. 1.1) освещалось интенсивным источником света. На основании принципа Гюйгенса это отверстие можно считать новым точечным источником полусферических волн. Эти волны падают на два малых отверстия S_1 и S_2 в следующем экране, которые в свою очередь становятся новыми точечными источниками волн.

Таким способом в опыте Юнга достигается разделение исходной волны на две. Эти волны налагаются друг на друга в области за отверстиями и могут интерферировать, так как источники S_1 и S_2 когерентны. На экране B образуется интерференционная картина.

Разделение волны от первичного некогерентного источника на две когерентные волны, т. е. получение двух вторичных когерентных точечных источников, может осуществляться разными способами. Но расчет интерференционной картины во всех таких случаях производится одинаково, так же, как и в схеме Юнга. Если в излучении первичного источника все независимые цуги волн характеризуются одной и той же длиной волны λ, то для излучения вторичных источников S_1 и S_2 можно использовать монохроматическую идеализацию, несмотря на то, что их излучение также представляет собой ту же хаотическую последовательность отдельных цугов. Замена такой последовательности цугов бесконечной синусоидальной волной возможна здесь потому, что точечные вторичные источники когерентны, а разность хода излучаемых ими волн в любой точке экрана B меньше протяженности отдельного цуга. Для этого, разумеется, экран B должен быть удален от источников S_1 и S_2 на значительное расстояние L, а расстояние d между источниками S_1 и S_2 должно быть достаточно мало.

Схема расчета интерференционной картины ясна из рис. 1.2. В точке O, расстояния до которой от источников S_1 и S_2 одинаковы, приходящие волны усиливают друг друга, так как колебания поля в этой точке происходят в одинаковой фазе. Результат сложения колебаний в произвольной точке P определяется разностью хода l волн, приходящих в P из S_1 и S_2. Если l равно целому числу длин волн λ, то колебания в P усиливают друг друга;

если l равно нечетному числу полуволн, то колебания взаимно ослабляются.

Выразим разность хода l волн, приходящих в точку P, через угол θ между осью и направлением на точку P и расстояние d между источниками (рис. 1.2).

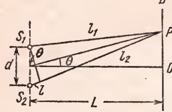


Рис. 1.2. К расчету интерференционной картины в опыте Юнга.

Будем считать, что $d \ll L$. Тогда при малых θ разность хода можно найти, опуская из S_1 перпендикуляр на прямую S_2P:

$$l = d\theta. \qquad (1.1)$$

Формула (1.1) дает возможность определить угловое положение максимумов и минимумов на экране B (рис. 1.2). Направления на максимумы получим, полагая в (1.1) $l = n\lambda$:

$$\theta_{max} = n\frac{\lambda}{d}, \qquad n = 0, \pm 1, \pm 2, \ldots \qquad (1.2)$$

Полагая $l = (2n+1)\lambda/2$, получим направления на минимумы:

$$\theta_{min} = \left(n + \frac{1}{2}\right)\frac{\lambda}{d}, \qquad n = 0, \pm 1, \pm 2, \ldots \qquad (1.3)$$

Угловое расстояние $\Delta\theta$ между соседними максимумами или минимумами, как видно из этих формул, равно λ/d, а расстояние h между ними на экране B, как видно из рис. 1.2, равно

$$h = L\,\Delta\theta = \frac{\lambda L}{d}. \qquad (1.4)$$

Мы рассмотрели только точки экрана B, лежащие в плоскости чертежа на рис. 1.2. Легко сообразить, что вся интерференционная картина на экране B в непосредственной близости от точки O состоит из светлых и темных полос, называемых интерференционными полосами. Они находятся на равных расстояниях друг от друга и направлены перпендикулярно плоскости чертежа, т. е. плоскости, в которой лежат первичный точечный источник S (т. е. отверстие A на рис. 1.1) и вторичные источники S_1 и S_2.

Трудность наблюдения интерференции света в таких опытах связана с тем, что длина волны света очень мала. При $\lambda=5\cdot10^{-5}$ см и расстоянии d между источниками, равном всего 0,5 мм, расстояние между соседними интерференционными полосами будет составлять только 1 мм при удалении экрана B на 1 м от источников.

Рассчитаем распределение освещенности на экране, где наблюдается интерференционная картина. Если при освещении экрана несколькими некогерентными источниками складываются освещенности, то при интерференции складываются напряженности полей в когерентных волнах, а освещенность в каждой точке экрана пропорциональна квадрату амплитуды результирующего колебания. Поскольку расстояния от источников до точки наблюдения почти одинаковы, то амплитуды обеих волн в точке наблюдения практически равны, если вторичные источники S_1 и S_2 одинаковы. При этом напряженность поля в точке наблюдения пропорциональна выражению

$$\cos\omega\left(t-\frac{l_1}{c}\right)+\cos\omega\left(t-\frac{l_2}{c}\right)= \\ =2\cos\frac{\omega}{2c}(l_2-l_1)\cos\omega\left(t-\frac{l_1+l_2}{2c}\right).$$

Видно, что это есть уравнение колебаний с амплитудой $2\cos(\omega l/2c)$, зависящей от разности хода $l=l_2-l_1$. Так как освещенность пропорциональна квадрату амплитуды, то распределение освещенности $E(\theta)$ в зависимости от направления на точку наблюдения θ имеет вид

$$E(\theta)\sim 4\cos^2\frac{\omega l}{2c}=4\cos^2\frac{\omega d}{2c}\theta=2\left(1+\cos\frac{\omega d}{c}\theta\right). \quad (1.5)$$

Распределение освещенности в интерференционных полосах, наблюдаемых по схеме Юнга, показано на рис. 1.3. Освещенность в минимумах равна нулю, а в максимумах она в четыре раза больше освещенности, которая создавалась бы одним источником. Положение максимумов и минимумов, как видно из формулы (1.5), находится в соответствии с полученными ранее формулами (1.2) и (1.3).

Отметим, что среднее значение освещенности по экрану в интерференционной картине, согласно формуле (1.5), равно просто удвоенной освещенности от одного источника.

Это значит, что при интерференции происходит только перераспределение энергии в пространстве.

Интенсивность наблюдаемой в опыте Юнга интерференционной картины можно заметно увеличить, если вместо точечных отверстий A, S_1 и S_2 в экранах использовать узкие, длинные, параллельные между собой щели. Вид полос вблизи центра интерференционного поля будет при этом таким же, как и при использовании точечных отверстий. Поясним это. Если точечное отверстие A перемещать перпендикулярно плоскости чертежа на рис. 1.1, то интерференционные полосы на экране, получаемые от точечных отверстий S_1 и S_2, просто будут смещаться вдоль своих направлений, т. е. также перпендикулярно плоскости чертежа. Поэтому замена отверстия A длинной щелью, т. е. непрерывной цепочкой точечных некогерентных источников, не приведет к ухудшению четкости интерференционных полос по крайней мере в той области, где их кривизна незначительна. Не приведет к ухудшению четкости и перемещение по этому направлению отверстий S_1 и S_2. Поэтому их также можно заменить на узкие длинные щели.

Рис. 1.3. Распределение освещенности в интерференционных полосах, наблюдаемых по схеме Юнга.

В качестве другого примера интерференционного опыта, который может быть рассчитан по схеме Юнга, рассмотрим метод наблюдения интерференции света с помощью зеркал Френеля. Вторичными когерентными источниками S_1 и S_2 здесь являются изображения реального источника света S в плоских зеркалах, расположенных под малым углом α друг к другу (рис. 1.4). Экран защищен от попадания прямых лучей от источника S.

Выясним прежде всего, в каких точках экрана будет наблюдаться интерференционная картина. Будет ли она занимать весь экран или какую-то его часть? Интерференция света наблюдается только там, где происходит наложение волн, пришедших от когерентных источников. В данном опыте нужно определить ту область экрана, в любую точку которой свет приходит после отражения от каждого из зеркал. На рис. 1.4 построены изображения S_1 и S_2 точечного источника S в обоих зеркалах. Нетрудно убедиться,

что центр окружности, на которой лежат точки S, S_1 и S_2, находится в точке A — вершине угла, образованного зеркалами. Из рисунка видно, что интерференционное поле на экране ограничено точками B_1 и B_2. Например, B_2 является крайней нижней точкой, в которую еще приходит свет после отражения от верхнего зеркала. Ее можно

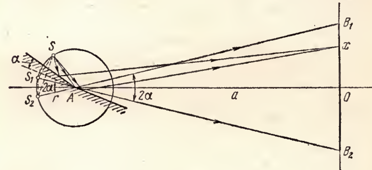


Рис. 1.4. Зеркала Френеля.

найти, если провести прямую через изображение источника S_1 в верхнем зеркале и точку A. Аналогично находится точка B_1. Из рис. 1.4 видно, что для расчета интерференционной картины действительно можно воспользоваться схемой Юнга и всеми полученными выше для нее формулами. Так как область интерференции на экране ограничена точками B_1 и B_2, то легко подсчитать число интерференционных полос. Если экран расположен так, как показано на рисунке, т. е. параллельно линии S_1S_2, соединяющей изображения источника в зеркалах, то $B_1B_2=2a\,\mathrm{tg}\,\alpha$, где a — расстояние от линии пересечения зеркал до экрана. Учитывая малость угла α между зеркалами, можно приближенно написать $B_1B_2\approx 2a\alpha$. Расстояние h между двумя соседними полосами дается формулой (1.4). Расстояние d между источниками S_1 и S_2, как видно из рис. 1.4, равно $2r\alpha$. Расстояние до экрана L равно $r+a$. Поэтому, согласно (1.4),

$$h=\frac{\lambda\,(r+a)}{2r\alpha}.$$

Полное число полос интерференции N, умещающихся на

интерфереиционоиом поле, равио

$$N=\frac{B_1B_2}{h}=\frac{4\alpha^2 ar}{\lambda(a+r)}.$$

Так как интерфереиционииую картииу удобио иаблюдать при условии $r \ll a$, то $N \approx 4\alpha^2 r/\lambda$.

Интерфереициоииые явлеиия, исторически послужившие эксперимеитальиым доказательством волиовой природы света, и в иаши дии иаходят важиые практические примеиеиия, в частиости в спектроскопии и в метрологии. Элемеитарная теория, осиоваииая иа использоваиии моиохроматической идеализации для световых воли, вполие пригодиа для описаиия действия приборов, примеияемых в большиистве интерфереициоииых исследоваиий. При этом можио считать, что поведеиие отдельиых пучков интерферирующих лучей подчииияется закоиам геометрической оптики, и полиостью преиебрегать дифракциоииыми явлеииями. Так мы и поступали при описаиии опыта с зеркалами Фреиеля.

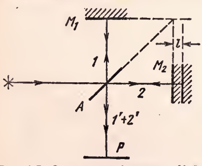


Рис. 1.5. Схема интерферометра Майкельсона.

В рассмотреииых выше интерфереициоииых опытах интерферируют волиы с малой разиостью хода — порядка иескольких длии волн. Но существуют и такие интерфереициоииые приборы, в которых разиость хода может быть сделаиа весьма большой. Эти приборы иазываются интерферометрами. На рис. 1.5 приведеиа упрощеииая схема интерферометра Майкельсоиа. Свет от иекоторого источиика падает слева на полупрозрачиое зеркало A и разделяется иа два пучка: отражеииый *1* и прошедший *2*. После отражеиия от зеркал M_1 и M_2 эти пучки снова падают иа полупрозрачное зеркало A и частичио отражаются, а частичио проходят через иего. В результате на экраие P можно иаблюдать интерфереицию лучей *1'* и *2'*. Картииа интерфереиции зависит от разиости хода лучей, которая

определяется разностью «плеч» интерферометра. Одно из зеркал (M_2 на рис. 1.5) может перемещаться с помощью микрометрического винта, оставаясь параллельным самому себе. При его перемещении изменяется разность хода и интерференционные полосы на экране P смещаются. Смещение интерференционной картины на одну полосу происходит при перемещении зеркала M_2 на расстояние, равное половине длины волны.

Интерферометр Майкельсона используется для выполнения особенно точных измерений длины. Например, в метрологии с его помощью производится сравнение первичного эталона длины, т. е. длины волны оранжевой линии криптона-86, со вторичными эталонами, выполненными в виде твердых стержней.

§ 2. Дифракция света

Характерной особенностью дифракционных явлений в оптике оказывается то, что здесь, как правило, длина волны света почти всегда много меньше размеров преград на пути световых волн. Поэтому наблюдать дифракцию света можно только на достаточно больших расстояниях от преграды. Проявление дифракции состоит в том, что распределение освещенности отличается от простой картины, предсказываемой геометрической оптикой на основе прямолинейного распространения света.

Строгий расчет дифракционной картины представляет собой очень сложную математическую задачу. Но в некоторых практически важных случаях достаточно хорошее приближение дает упрощенный подход, основанный на использовании принципа Гюйгенса — Френеля.

Пусть поверхность S представляет собой положение волновой поверхности в некоторый момент времени (рис. 2.1). Для того чтобы определить вызванные волной колебания в некоторой точке P, нужно, по Френелю, определить колебания, вызываемые в этой точке отдельными вторичными волнами, приходящими в нее от отдельных элементов поверхности S, и затем сложить эти колебания с учетом их амплитуд и фаз. При этом следует считать, что в точке P сказывается влияние только той части волновой поверхности S, которая не загораживается каким-либо препятствием.

Проиллюстрируем применение принципа Гюйгенса — Френеля на следующем примере. Пусть на непрозрачную преграду с круглым отверстием падает слева плоская монохроматическая волна (рис. 2.2). Такую волну можно

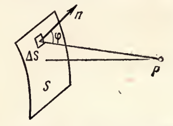


Рис. 2.1. К расчету дифракции на основе принципа Гюйгенса — Френеля.

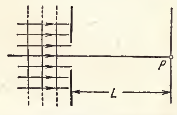


Рис. 2.2. Падение плоской монохроматической волны на преграду с круглым отверстием.

получить, например, от точечного источника монохроматического света, удаленного на бесконечность или помещенного в фокус собирающей линзы большого диаметра. Будем интересоваться освещенностью экрана в точке P, находящейся на оси симметрии. Для учета интерференции вторичных волн Френель предложил мысленно разбить волновую поверхность падающей волны в месте расположения преграды на кольцевые зоны (зоны Френеля) по следующему правилу: расстояния от краев соседних зон до точки P (рис. 2.3) должны отличаться на половину длины волны, т. е.

$$l_1=L+\frac{\lambda}{2},\quad l_2=L+2\frac{\lambda}{2},\quad \ldots,\quad l_k=L+k\frac{\lambda}{2}. \quad (2.1)$$

Если смотреть на волновую поверхность из точки P, то зоны Френеля будут выглядеть так, как показано на рис. 2.4. Из рис. 2.3 легко найти радиусы зон Френеля:

$$r_k=\sqrt{l_k^2-L^2}=\sqrt{k\lambda L+k^2\frac{\lambda^2}{4}}\approx\sqrt{k\lambda L}. \quad (2.2)$$

Видно, что радиус k-й зоны пропорционален $\sqrt{k}$, если $k\lambda/L\ll 1$. При выполнении этого условия площади зон Френеля можно считать одинаковыми. Результат интерференции вторичных волн в точке P, как мы увидим ниже,

определяется тем, сколько зон Френеля открывает круглое отверстие на волновой поверхности.

Предположим, что отверстие в преграде представляет собой диафрагму, диаметр которой можно изменять. Пусть сначала радиус отверстия много меньше радиуса первой зоны Френеля. Тогда можно считать, что колебания от

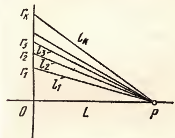


Рис. 2.3. Построение зон Френеля.

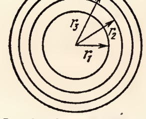


Рис. 2.4. Зоны Френеля.

всех точек волновой поверхности в этом маленьком отверстии приходят в точку P практически в одинаковой фазе. Изобразим колебание поля в точке P, вызванное этой вторичной волной, с помощью векторной диаграммы (рис. 2.5, a). Этому колебанию на ней сопоставляется вектор ΔA_1, который вращается с угловой скоростью ω, равной циклической частоте падающей волны, в направлении против часовой стрелки. Увеличим отверстие диафрагмы еще немного, так чтобы площадь его удвоилась. Колебания, приходящие в точку P от вновь открытого участка волновой поверхности, несколько отстают по фазе и изображаются на диаграмме вектором ΔA_2. Длина этого вектора равна длине вектора ΔA_1, так как равны между собой площади соответствующих им участков волновой поверхности. Продолжая увеличивать отверстие диафрагмы, будем откладывать на диаграмме векторы, соответствующие приходящим в точку P колебаниям от вновь открываемых участков волновой поверхности. Колебаниям, приходящим в P от участка, прилегающего к границе первой зоны Френеля, будет соответствовать вектор ΔA_n, повернутый относительно ΔA_1 на π, так как, согласно определению зон Френеля, разность хода соответствующих им вторичных волн равна $\lambda/2$.

Результнрующее колебанне в точке P, создаваемое волиой, которая прошла через круглое отверстие, совпадающее с первой зоной Френеля, изображается вектором A_1, равным сумме векторов ΔA_i (рис. 2.5, *а*). Будем увеличивать отверстие диафрагмы дальше. Когда иа ием будут умещаться две первые зоны Френеля, векториая диаграмма колебаинй в точке P примет вид, изображеииый иа рис. 2.5,*б*. При строгом равеистве амплитуд складываемых колебаиий ΔA_i амплитуда результирующего колебания A_2 должиа была бы равияться нулю, т. е. вторичиые волиы прн двух открытых зонах Френеля полиостью гасилн бы друг друга в точке P. Одиако действие даже одииаковых по площади участков волиовой поверхиости в точке P иесколько убывает по мере увеличеиия угла φ между иаправлеиием на точку P н нормалью к волиовой поверхности (рис. 2.1). Поэтому в действительиости амплитуда A_2 нмеет коиечное, хотя и очеиь малое значеиие.

Рис. 2.5. Расчет амплитуды результирующего колебания в точке P с помощью векториых диаграмм: *а*) в отверстии укладывается одиа зона Френеля, *б*) две зоиы Френеля.

Таким образом, освещеииость экраиа в точке P, пропорцнональиая квадрату амплитуды результирующего колебаиия, будет по мере увеличеиия отверстия круглой диафрагмы меияться не моиотоино. Пока открывается первая зоиа Френеля, освещенность в P увеличивается и стаиовится максимальной прн полиостью открытой первой зоие. По мере открываиня второй зоиы Фреиеля освещеиность убывает и при полностью открытой второй зоне умеиьшается почти до иуля. Затем освещенность будет увеличиваться сиова, и т. д. Этн иа первый взгляд парадоксальиые результаты, предсказываемые иа основе прниципа Гюйгеиса — Френеля, иаходятся в хорошем согласни с экспериметом. Подчеркием, что они иаходятся в вопиющем протнворечии с предсказаииямн геометрической оптики, согласио которой при падеиии плоской волиы освещеиность в точке P, лежащей на осн круглого отверстня, не зависнт от диаметра отверстия.

Наиболее неожиданным в полученных выше результатах является, пожалуй, то, что при двух открытых зонах Френеля (и вообще при небольшом четном числе открытых зон) освещенность в точке P близка к нулю. Не менее неожиданным является то, что в точке P позади непрозрачного круглого экрана, расположенного на месте преграды с отверстием, освещенность не будет равна нулю, как это следовало бы из геометрической оптики. Если при этом непрозрачный круглый экран перекрывает лишь несколько первых зон Френеля, то в точке P освещенность будет почти такой же, как и без экрана. В этом можно убедиться, если рассматривать вектор A, изображающий колебания напряженности поля в точке P при полностью открытой волновой поверхности, как сумму двух векторов, один из которых изображает колебания от открытого участка волновой поверхности, а другой — от тех зон Френеля, которые перекрыты экраном. В центре геометрической тени оказывается свет.

Теперь не представляет труда оценить те условия наблюдения, при которых дифракционные явления становятся существенными и картина распределения освещенности на экране заметно отличается от предсказываемой геометрической оптикой. По геометрической оптике распределение освещенности на экране должно соответствовать форме отверстия, так что освещенность экрана равна нулю в области геометрической тени, а в точке P такая же, как и в отсутствие преграды. Но мы видели, что в случае, когда на отверстии укладывается лишь несколько зон Френеля, освещенность в точке P совсем иная. Это дает возможность оценить то расстояние L от отверстия до точки наблюдения, на котором именно дифракционные явления определяют наблюдаемую картину. Для этого в формуле (2.2) следует считать $k \sim 1$, а r_k положить равным размеру отверстия (или преграды) d.

В результате находим

$$L \sim \frac{d^2}{\lambda}. \tag{2.3}$$

Построения Френеля позволяют легко рассчитать освещенность позади непрозрачного круглого экрана или экрана с круглым отверстием только в точках, лежащих

на оси симметрии. Найти вид всей дифракционной картины на экране очень трудно.

Но можно осуществить такие условия наблюдения дифракции света, при которых возможен полный расчет распределения освещенности в дифракционной картине на экране. Пусть плоская монохроматическая волна от бесконечно удаленного точечного источника падает на экран S с отверстием, а дифракционная картина наблюдается на экране в фокальной плоскости линзы (рис. 2.6). Так как в каждой точке фокальной плоскости линзы, например P на рис. 2.6, сходятся лучи, которые до линзы были параллельны между собой, то наблюдаемая здесь картина называется дифракцией в параллельных лучах. Так как линза не вносит дополнительной разности хода между параллельными до линзы лучами, то складывающиеся в точке P колебания имеют такую же разность фаз, как и до линзы на плоскости, перпендикулярной к этим лучам. Такая схема наблюдения дифракции была предложена Фраунгофером.

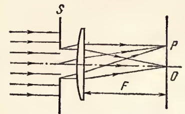


Рис. 2.6. Наблюдение дифракции в параллельных лучах.

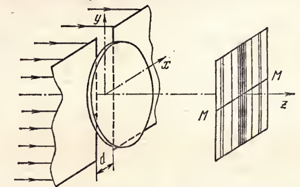


Рис. 2.7. Наблюдение дифракции от щели с параллельными краями.

Пусть отверстие в экране S представляет собой щель шириной d (рис. 2.7), которую считаем бесконечно протя-

женной в направлении оси y. Построенные по принципу Гюйгенса волновые поверхности позади щели представляют собой цилиндрические поверхности с образующей, параллельной краям щели (рис. 2.8). Так как волновая поверхность в направлении оси y не ограничена, то дифракционных эффектов в этом направлении быть не может. Поэтому весь прошедший через линзу и попадающий на экран дифрагированный свет будет сосредоточен вдоль линии MM, лежащей в плоскости xz. Вместо изображения точечного источника в фокальной плоскости линзы, которое было бы в отсутствие щели, получается дифракционная картина, вытянутая вдоль линии MM.

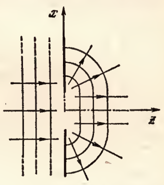


Рис. 2.8. Волновые поверхности, построенные по принципу Гюйгенса.

Если создающий падающую волну точечный источник сместить вдоль оси y так, чтобы падающие на щель параллельные лучи образовали некоторый угол с осью z, то дифракционная картина на экране, не изменяя своего вида, сместится из положения MM на такой же угол. Поэтому при замене точечного источника света на тонкую светящуюся линию, параллельную оси y, каждый ее точечный элемент будет создавать свою дифракционную картину, параллельную MM, а вся дифракционная картина на экране будет состоять из параллельных светлых и темных полос, как показано на рис. 2.7. Для ее нахождения достаточно рассмотреть только плоскость xz.

Согласно принципу Гюйгенса — Френеля волновую поверхность падающей волны в щели на оси x следует разбить на столь малые участки, чтобы колебания в точке наблюдения P, вызываемые вторичными волнами от всех точек одного участка, имели бы почти одинаковую фазу. Колебания в P, вызываемые вторичными волнами, распространяющимися под углом θ от разных участков (рис. 2.9), следует просуммировать с учетом сдвигов по фазе. Это удобно сделать с помощью векторной диаграммы, построенной на рис. 2.10. Вектор ΔA_1 изображает колебание, при-

ходящее в точку P от участка Δx_1, лежащего вблизи нижнего края щели. Вектор ΔA_2, изображающий колебание от соседнего участка Δx_2, повернут относительно ΔA_1 на некоторый небольшой угол. Вектор ΔA_n, изображающий колебание от последнего участка Δx_n, лежащего у верхнего края щели, повернут относительно вектора ΔA_1 на

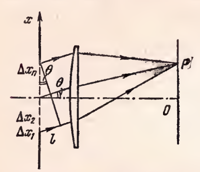


Рис. 2.9. К расчету суммарного колебания в точке P.

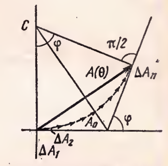


Рис. 2.10. Сложение колебаний в точке P с помощью векторной диаграммы.

угол φ, соответствующий разности хода $l=d\sin\theta$ (рис. 2.9) между лучами, приходящими от краев щели. Чтобы найти сдвиг по фазе φ между колебаниями в P, вызванными волнами с разностью хода l, следует учесть, что сдвиг по фазе равен 2π при разности хода λ:

$$\varphi = 2\pi\frac{l}{\lambda} = 2\pi\frac{d\sin\theta}{\lambda}. \qquad (2.4)$$

Найдем длину суммарного вектора $A(\theta)$, которая равна амплитуде колебаний в точке наблюдения P. Легко видеть, что вектор $A(\theta)$ представляет собой хорду окружности с центром в точке C (рис. 2.10). Прежде всего отметим, что длина дуги, стягиваемой хордой $A(\theta)$, равна амплитуде колебаний A_0 в точке O на экране, так как в эту точку вторичные волны от всех участков Δx_i, распространяясь под углом $\theta=0$, приходят в одинаковой фазе и все векторы ΔA_n имеют для точки O одинаковые направления. Длину дуги A_0 и длину хорды $A(\theta)$ легко связать между собой из геометрических соображений.

Из рис. 2.10 видно, что

$$\frac{A(\theta)}{2} = R\sin\frac{\varphi}{2} = \frac{A_0}{\varphi}\sin\frac{\varphi}{2},$$

откуда

$$A(\theta) = A_0 \frac{\sin(\varphi/2)}{\varphi/2}. \qquad (2.5)$$

Освещениость экрана $E(\theta)$ в точке P, пропорциоиальная квадрату амплитуды колебаний, связана с освещенностью E_0 в точке O, согласио (2.5), следующим соотношеиием:

$$E(\theta) = E_0 \frac{\sin^2(\varphi/2)}{(\varphi/2)^2}, \qquad (2.6)$$

где φ дается формулой (2.4). Распределеиие освещеииости на экраие $E(\theta)$ при дифракции плоской волны иа длииной щели показаио на рис. 2.11. Вместо бесконечио узкой лииии, которая получалась бы в фокальной плоскости линзы согласио законам геометрической оптики, иа экраие получаются дифракциоииые полосы, параллельные щели. Рядом с яркой центральиой полосой будут слабые побочиые полосы, отделеиные друг от друга полной темиотой, причем ширииа побочиых полос вдвое меньше ширины центральиой. Освещеииость в центре первой побочиой полосы, как видно из формулы (2.6), почти в 25 раз меиьше освещеииости в цеитре картииы. Освещеииость обращается в иуль тогда, когда аргумеит синуса в (2.6) кратеи π. Это соответствует углам дифракции θ, при которых, как видио из (2.4),

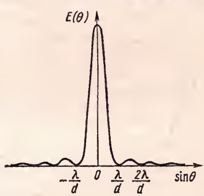


Рис. 2.11. Распределение освещенности на экране при дифракции плоской волны иа щели.

$$d\sin\theta = k\lambda, \qquad k = \pm 1,\ \pm 2, \ldots \qquad (2.7)$$

Отметим, что положеиие мииимумов освещеииости легко найти и без помощи формулы (2.6). Для этого достаточио только сообразить, что минимумам соответствует разность

хода l между крайними лучами (рис. 2.9), равная целому числу длин волн λ. Действительно, если разность хода l равна, например, λ, то всю щель можно разбить на пары одинаковых участков, отстоящих друг от друга на $d/2$. Разность хода вторичных волн от каждой такой пары равна $\lambda/2$, и эти волны в точке наблюдения гасят друг друга.

Чем у́же щель, тем шире дифракционные полосы. Из формулы (2.7) видно, что при уменьшении ширины щели d до размеров порядка длины волны λ центральная полоса расплывается на весь экран.

§ 3. Спектральные приборы. Дифракционная решетка

Назначение спектральных приборов — исследовать спектральный состав излучения, т. е. определять, из каких монохроматических волн оно состоит. Иначе говоря, спектральный прибор производит гармонический анализ излучения. Действие спектральных приборов основано на том,

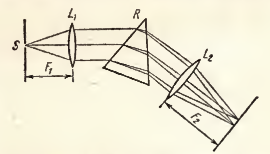


Рис. 3.1. Принципиальная схема спектрального прибора.

что в некоторых физических системах условия прохождения света разной длины волны оказываются различными. Такие системы носят название диспергирующих.

В экспериментах по изучению спектров обычно используют призму или дифракционную решетку. Принципиальная схема простейшего спектрального прибора показана на рис. 3.1. Щель S, на которую падает исследуемое излучение, находится в фокальной плоскости линзы L_1. Эта часть прибора называется коллиматором. Выходящий из линзы параллельный пучок света падает на призму R. Вследствие дисперсии света в веществе призмы свет разных

длин волн выходит из призмы под разными углами. В фокальной плоскости линзы L_2 помещается экран или фотопластинка, на которой фиксируется приходящее излучение. Линза фокусирует параллельные пучки лучей, и в результате образуются изображения входной щели в разных местах экрана для разных длин волн.

Идеальным был бы такой спектральный прибор, распределение энергии падающего излучения на выходе которого определялось бы только спектральным составом излучения и не зависело бы от конструкции прибора. Но любой реальный спектральный прибор всегда вносит искажения. Идеальный прибор при падении монохроматического излучения давал бы на выходе единственную бесконечно узкую спектральную линию. Однако в реальном приборе на выходе вместо узкой линии получается некоторое распределение освещенности, характеризуемое контуром определенной формы. Этот контур имеет конечную ширину, что ограничивает способность прибора разделять две близко расположенные спектральные линии.

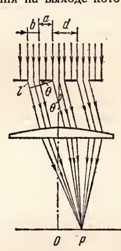


Рис. 3.2. К объяснению действия дифракционной решетки.

В наиболее совершенных спектральных приборах в качестве диспергирующего элемента используются дифракционные решетки. Лучшие дифракционные решетки представляют собой полированные стеклянные или металлические пластины, на которых алмазным резцом нанесены при помощи специальной делительной машины параллельные одинаковые штрихи, расположенные на строго одинаковых расстояниях друг от друга.

Действие дифракционной решетки можно понять, рассматривая падение плоской монохроматической волны на регулярную периодическую структуру, состоящую из чередующихся параллельных друг другу щелей одинаковой ширины b, расположенных на одинаковом расстоянии a друг от друга (рис. 3.2). Сумма $a+b$ является периодом этой структуры и называется постоянной решетки d.

В каждой точке P на экране в фокальной плоскости линзы соберутся те лучи, которые до линзы были параллельны между собой и распространялись под определенным углом θ к направлению падающей волны. Поэтому освещенность в точке P определяется результатом интерференции вторичных волн, распространяющихся как от разных участков одной щели, так и от разных щелей. Колебание в точке P, вызываемое вторичными волнами от одной щели, было рассмотрено в предыдущем параграфе. Поэтому мы можем считать это колебание известным и для нахождения результирующего колебания сложить колебания от всех щелей с учетом сдвига по фазе между ними.

Легко найти те направления, распространяясь по которым вторичные волны от всех щелей будут приходить в точку P в фазе и усиливать друг друга. Так будет, если разность хода l между вторичными волнами, идущими из эквивалентных точек соседних щелей, равна целому числу длин волн (рис. 3.2):

$$d\sin\theta = n\lambda, \qquad n = 0,\ \pm 1,\ \pm 2,\ \ldots \tag{3.1}$$

В точках на экране, где собираются лучи, распространяющиеся под углами θ_n, удовлетворяющими условию (3.1), расположены так называемые главные максимумы дифракционной картины. А какой вид имеет дифракционная картина между главными максимумами? Чтобы выяснить это, возьмем определенную решетку, имеющую большое число N периодов, и рассмотрим, как будет меняться освещенность на экране при постепенном переходе от главного максимума нулевого порядка ($n=0$) к главному максимуму первого порядка ($n=1$). Для нахождения амплитуды результирующего колебания воспользуемся методом векторных диаграмм. При $\theta=0$ векторы A_i, изображающие колебания от разных щелей, параллельны друг другу и при сложении дадут вектор $A(0)$, длина которого A_0 в N раз больше длины вектора A_i и равна амплитуде колебаний в главном максимуме. По мере увеличения угла θ между колебаниями от соседних щелей появляется одинаковый сдвиг по фазе ψ, так что одинаковые по величине соседние векторы A_i повернуты друг относительно друга на угол ψ. В результате получается ломаная линия, вписанная в окружность (рис. 3.3, *а*). Если эта ломаная линия окажется замкнутой, то амплитуда результирующего колебания $A(\theta)$ обратится

в нуль, что приводит к полной темноте в соответствующих точках экрана. Замыкание ломаной линии из векторов A_l происходит при выполнении условия

$$N\psi = 2k\pi, \qquad k = 1, 2, \ldots, N-1. \tag{3.2}$$

Значение $k=1$ соответствует разности хода между пучками от первой и последней (т. е. N-й) щели, равной одной длине волны λ. Значение $k=2$ соответствует разности хода, равной 2λ, и т. д. Значение $k=N$ в формуле (3.2) соответствовало бы разности хода между двумя пучками от соседних щелей, равной длине волны λ. Но это, как видно из формулы (3.1), есть как раз условие главного максимума первого порядка. Поэтому между главными максимумами нулевого и первого порядков располагаются $N-1$ минимумов. Угловое положение этих минимумов определяется из соотношения (3.2) при учете, что сдвиг по фазе ψ между пучками от двух соседних щелей выражается через разность хода l между ними следующим образом (рис. 3.2):

$$\psi = 2\pi\frac{l}{\lambda} = 2\pi\frac{d\sin\theta}{\lambda}. \tag{3.3}$$

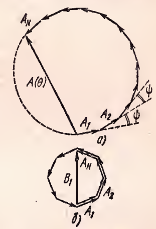


Рис. 3.3. Векторные диаграммы для нахождения результирующего колебания в точке P. Случай *б* соответствует первому побочному максимуму.

Подставляя (3.3) в формулу (3.2), получаем условие для нахождения направлений на минимумы $\theta_{\min}$:

$$Nd\sin\theta_{\min} = k\lambda, \qquad k = 1, 2, \ldots, N-1. \tag{3.4}$$

Положения $N-1$ минимумов между главными максимумами первого и второго порядков даются той же формулой (3.4), в которой k уже пробегает значения от $N+1$ до $2N-1$, и т. д. Очевидно, что между $N-1$ минимумами располагаются $N-2$ максимумов, которые в отличие от главных называются побочными. Эти максимумы возникают, когда ломаная линия на векторной диаграмме, образованная векторами A_l, частично налагаясь сама на себя, оканчи-

вается в верхней точке окружности, так что замыкающий ее вектор результирующего колебания $A(\theta)$ проходит по диаметру окружности. На рис. 3.3, *б* показана векторная диаграмма, соответствующая направлению на первый побочный максимум, расположенный рядом с главным. С помощью рисунка легко видеть, что при большом числе штрихов N амплитуда колебаний в этом максимуме B_1 связана

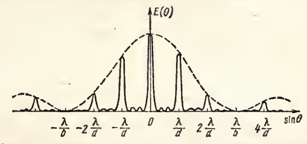


Рис. 3.4. Распределение освещенности для дифракционной решетки, содержащей четыре щели.

с амплитудой колебаний в главном максимуме A_0 соотношением $B_1=\frac{2}{3\pi}A_0$. Пропорциональная квадрату амплитуды освещенность будет в центре первого побочного максимума почти в 25 раз меньше, чем в центре главного. Остальные побочные максимумы будут еще слабее.

Как мы видели, амплитуда колебаний в главном максимуме в N раз больше амплитуды колебаний, создаваемых в этой точке вторичной волной от одной щели. Но амплитуда колебаний от одной щели зависит от угла θ и определяется формулой (2.5) предыдущего параграфа. Поэтому амплитуды результирующего колебания в главных максимумах разных порядков различны. Картина распределения освещенности на экране показана на рис. 3.4 для дифракционной решетки, содержащей $N=4$ щели, причем ширина щели b в три раза меньше периода d. Огибающая главных максимумов соответствует распределению освещенности в дифракционной картине от одной щели (рис. 2.11).

Легко видеть, что главный максимум определенного порядка пропадает, если его положение совпадает с каким-нибудь минимумом дифракционной картины от одной щели. Сравнивая условия $d\sin\theta=n\lambda$, определяющие глав-

ный максимум решетки, с условием минимума в дифракционной картине от одной щели $b \sin\theta = k\lambda$, видим, что условие исчезновения главного максимума n-го порядка можно записать в виде

$$\frac{b}{d} = \frac{k}{n}, \qquad k = 1,\ 2,\ \ldots,\ n-1. \tag{3.5}$$

Здесь k не может принимать значение, равное n, ибо при этом решетки уже нет. На рис. 3.4 $b/d = 1/3$, поэтому отсутствует главный максимум третьего порядка. Таким образом, распределение энергии падающей на решетку плоской монохроматической волны по главным дифракционным максимумам разных порядков зависит от отношения b/d, а в общем случае определяется структурой одного периода решетки.

Положение главных дифракционных максимумов, определяемое формулой (3.1), при данной длине волны λ зависит только от периода решетки d. Оно не зависит ни от полного числа штрихов решетки N, ни от структуры каждого отдельного периода решетки. При увеличении полного числа штрихов N главные максимумы, оставаясь на прежних местах, становятся все резче и резче, так как между ними появляется все большее и большее число примерно равноотстоящих побочных максимумов.

Мы видим, что при использовании дифракционной решетки в качестве диспергирующего элемента спектрального прибора при падении монохроматической волны получается не одна спектральная линия, а набор главных максимумов конечной ширины. Если падающее излучение содержит свет нескольких длин волн λ_1, λ_2, ..., то главный максимум нулевого порядка для всех λ будет в одном и том же месте при $\theta = 0$, а положение главных максимумов первого, второго и т. д. порядков для разных длин волн будет различным в соответствии с формулой (3.1). Поэтому различают создаваемые решеткой спектры первого, второго и более высоких порядков.

Одной из важнейших характеристик дифракционной решетки является ее разрешающая способность, которая характеризует возможность разделить в падающем излучении две близкие длины волны λ и $\lambda + \Delta\lambda$. Разрешающей способностью называется отношение λ к минимально возможному значению $\Delta\lambda$, т. е. $\lambda/\Delta\lambda$. Считается, что две линии

спектра, создаваемого решеткой, различимы, если главный максимум n-го порядка для длины волны $\lambda+\Delta\lambda$ подходит к n-му главному максимуму для длины волны λ не ближе, чем ближайший минимум для λ (рис. 3.5). Этот условный критерий разрешимости спектральных линий был предложен Релеем. При выполнении критерия Релея налагающиеся дифракционные картины образуют максимум с небольшим провалом посредине (рис. 3.5), что воспринимается глазом по контрасту как наличие темного промежутка между максимумами для λ и $\lambda+\Delta\lambda$.

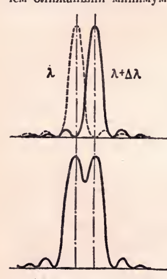


Рис. 3.5. Критерий Релея разрешимости спектральных линий.

Положение ближайшего к n-му главному максимуму минимума для длины волны λ определяется, в соответствии с формулой (3.4), следующим соотношением:

$$Nd\sin\theta=(Nn+1)\lambda. \quad (3.6)$$

Для положения n-го главного максимума для волны $\lambda+\Delta\lambda$, согласно формуле (3.1), можно написать

$$Nd\sin\theta=Nn(\lambda+\Delta\lambda). \quad (3.7)$$

При выполнении критерия Релея левые части (3.6) и (3.7) совпадают. Поэтому

$$(Nn+1)\lambda=Nn(\lambda+\Delta\lambda),$$

откуда

$$\frac{\lambda}{\Delta\lambda}=nN. \quad (3.8)$$

Разрешающая способность дифракционной решетки тем выше, чем больше в ней штрихов N и чем выше используемый порядок спектра n. Наибольший порядок спектра n_{max}, который можно получить с данной решеткой, ограничивается условием $|\sin\theta|\leqslant 1$:

$$d=n_{max}\lambda.$$

Чтобы решетка давала спектр хотя бы первого порядка, необходимо, чтобы период решеткн d был ие меньше длины волны λ.

Современиые решетки имеют до 1200 штрихов на 1 мм и разрешающую способиость в спектре первого порядка до 100 000.

В заключение вернемся к опыту Юига, рассмотрениому в § 1. Напомним, что этот опыт был рассмотрен в приближенни точечных вторнчных источников, которое применнмо тогда, когда размер отверстий миого меньше длины

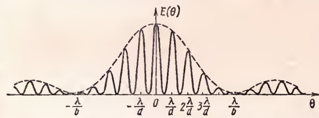


Рис. 3.6. Распределеиие освещенности в опыте Юига в случае, когда шириия щели в 5 раз меиьше расстояиия между цеитрами щелей. Пуиктириая огибающая соответствует дифракционной картиие от одиой щели.

волиы. Как мы виделн, вместо точечных отверстий можно использовать бесконечно узкие параллельные щелн и внд интерференцноиной картниы в центре экрана прн этом не меняется.

Одиако в реальном опыте щели всегда имеют конечную шнрнну. Теперь, после того как мы познакомнлнсь с действнем дифракционной решетки, легко выяснить, как изменяется интерференцнонная картниа в опыте Юига прн переходе к щелям коиечиой ширнны. Для этого достаточно сообразить, что щели в опыте Юига представляют собой дифракцнонную решетку, у которой полное чнсло штрихов N равио двум. Правда, в опыте Юнга отсутствует лннза, но иаблюдаемая иа удалеином экране ннтерференцнонная картииа практическн не отлнчается от той, которая иаблюдается в фокальиой плоскости линзы для решеткн с двумя щелямн, ибо прнходящне в одну и ту же точку удалениого экрана лучи от двух близкнх щелей почти параллельиы. На рис. 3.6 показано распределение освещеииости иа

удаленном экране в опыте Юнга в случае, когда ширина щели b в 5 раз меньше расстояния между центрами щелей d. Пунктир соответствует распределению освещенности экрана от одной щели.

§ 4. Протяженные источники света. Звездный интерферометр

Все рассмотренные выше интерференционные и дифракционные явления волновой оптики относились к случаю монохроматического света, излучаемого точечным источником. Однако все реальные источники света имеют конечные размеры, а излучаемый ими свет, как мы уже обсуждали выше, никогда не является строго монохроматическим. Поэтому интересно выяснить, к каким изменениям в результатах приведет отказ от монохроматической идеализации и учет конечных размеров источников света. Для простоты и большей наглядности выясним роль каждого из этих факторов в отдельности.

Начнем с учета конечных размеров источника. Будем считать, что реальный протяженный источник состоит из большого числа точечных взаимно некогерентных элементов, излучающих свет определенной длины волны. В этом случае интенсивность в любой точке волнового поля равна сумме интенсивностей от каждого точечного источника.

Рассмотрим изменение интерференционной картины в опыте Юнга, обусловленное использованием протяженного источника света. Разумеется, речь идет не об увеличении размеров источника в направлении, параллельном щелям: при использовании такого линейного источника вид интерференционной картины, как мы видели, не меняется. Речь идет и не об увеличении размеров вторичных когерентных источников, т. е. ширины щелей при использовании точечного первичного источника света,— этот случай уже был рассмотрен в конце предыдущего параграфа. Сейчас нас будет интересовать вид интерференционной картины при использовании первичного источника конечной ширины, а сами щели будем для простоты считать бесконечно узкими. Мы увидим, что с увеличением ширины источника резкость интерференционных полос уменьшается вплоть до их полного исчезновения. Это накладывает определенные условия

на размеры источников света в направлении, соединяющем отверстия или щели, при их использовании в интерференционных экспериментах по схеме Юнга.

Явление уменьшения резкости интерференционных полос, с которым приходится бороться в лабораторных экспериментах, нашло совершенно неожиданное и очень эффективное применение в астрономии.

Одной из важнейших астрономических задач является определение углового расстояния двойных звезд, т. е. того угла, под которым видны эти звезды с Земли. Если звезды находятся на очень маленьком угловом расстоянии θ друг от друга, то даже с помощью самых совершенных телескопов эту задачу решить не удается, так как в фокальной плоскости телескопа изображения этих звезд размыты вследствие явления дифракции и не могут быть разрешены.

Выясним прежде всего, как выглядит в фокальной плоскости объектива телескопа изображение звезды, которую из-за очень большого удаления можно считать точечным источником. Чтобы получить представление об этом, будем пока считать, что перед объективом телескопа помещена длинная щель шириной d с параллельными прямыми краями. Поскольку приходящий от звезды свет можно рассматривать как плоскую волну, в фокальной плоскости объектива будет наблюдаться дифракционная картина от щели, которая была описана при рассмотрении дифракции в параллельных лучах. Распределение освещенности для этого случая показано на рис. 2.11.

Освещенность экрана в первом боковом максимуме составляет, как было показано, менее 5% от освещенности в центре дифракционной картины. Это означает, что почти весь поток света, прошедший через щель, распространяется в интервале углов от $-\theta_1$ до θ_1, где угол θ_1 определяется формулой (2.7) при $k=1$. Поскольку мы здесь рассматриваем щель, ширина которой d много больше длины волны λ, то $\sin\theta_1$ можно заменить на θ_1, и тогда

$$\theta_1=\frac{\lambda}{d}. \qquad (4.1)$$

Изображение бесконечно удаленного точечного источника в фокальной плоскости линзы оказывается размытым в полоску, перпендикулярную краям щели. Длина a этой полоски ограничена размером центрального

дифракционного максимума и поэтому равна

$$a = 2F\theta_1 = 2\frac{\lambda}{d}F, \qquad (4.2)$$

где F — фокусное расстояние линзы.

Рассмотрение дифракции на круглом отверстии диаметром D показывает, что изображение бесконечно удаленного точечного источника размывается в круглое пятно, диаметр которого определяется той же формулой (4.2) с дополнительным числовым коэффициентом, близким к единице.

В телескопе роль круглого отверстия играет оправа объектива, и изображение звезды представляет собой создаваемую этой оправой дифракционную картину. Объективы большого диаметра могут давать изображение более высокого качества, так как, как видно из формулы (4.2), уменьшается влияние дифракции.

Если в телескоп наблюдают две звезды, находящиеся на малом угловом расстоянии друг от друга, то дифракционные картины, создаваемые каждой звездой, налагаются одна на другую. Если при этом главные максимумы дифракционных картин сближаются на расстояние, меньшее радиуса центрального дифракционного пятна, то, согласно критерию Релея, измерить точно расстояние между ними, а тем самым и угловое расстояние между звездами, невозможно. Правда, современные методы обработки экспериментальных результатов позволяют разрешать дифракционные картины, для которых критерий Релея, строго говоря, не выполняется. Однако для оценки разрешающей способности телескопа это не принципиально.

Итак, минимальное угловое расстояние θ между звездами, которое можно надежно измерить с помощью телескопа, равно

$$\theta \approx \theta_1 = \frac{\lambda}{D}. \qquad (4.3)$$

По этой формуле легко оценить, что на крупнейшем в мире телескопе — рефлекторе с диаметром зеркала $D = 6$ м, который построен в СССР,— можно измерять угловые размеры, не меньшие 0,02″.

Для измерения еще меньших угловых размеров используется звездный интерферометр Майкельсона, идею которого можно понять из рис. 4.1. Основными элементами ин-

терферометра являются непрозрачный экран A с двумя отверстнями, расстояние между которыми d можно изменять, собирающая линза L, расположенная непосредственно за экраном A, и экран B, находящийся в фокальной плоскости линзы, на котором наблюдаются ннтерференционные полосы.

Это есть одна нз возможных реализаций опыта Юнга, отличающаяся от рассмотренной ранее тем, что ннтерференционные полосы наблюдаются не на удаленном экране, а в фокальной плоскостн линзы, роль которой в звездном интерферометре выполняет объектнв телескопа-рефрактора или зеркало телескопа-рефлектора. При изменении расстояния d между отверстиями резкость ннтерференционных полос изменяется, и по этнм нзмененням можно определнть угловой размер двойной звезды θ.

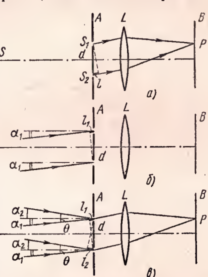


Рис. 4.1. К объяснеиию принципа действия звездиого интерферометра.

Чтобы понять, почему меняется резкость полос, рассмотрим сначала интерференционную картину, создаваемую одним бесконечно удаленным точечным источником S. Если нсточник S расположен на оптнческой оси (рнс. 4.1, a), то фазы вторичных нсточннков S_1 н S_2 совпадают и в некоторой точке P на экране B будет находиться светлая нли темная полоса в зависимостн от того, будет ли разность хода лучей l равна четному нли нечетному числу полуволн. Если источник S смещен с оптической осн прибора на угол α_1, то создаваемая им интерференционная картина окажется сдвинутой, поскольку между вторичными источниками имеется разность фаз, обусловленная разностью хода лучей l_1 от источника S до отверстий в экране A

(рис. 4.1, *б*):

$$l_1 = d\alpha_1. \qquad (4.4)$$

При наблюдении двойной звезды, которую можно рассматривать как два взаимно некогерентных точечных источника, на экране B будут налагаться две независимые интерференционные картины, создаваемые каждой звездой, и освещенность в любой точке экрана будет равна сумме освещенностей от каждой интерференционной картины. Как будет выглядеть эта суммарная интерференционная картина? Она будет отчетливой, если светлые полосы одной картины приходятся на светлые полосы другой, и исчезнет совсем, если светлые полосы одной совпадут с темными полосами другой. Полное исчезновение полос произойдет, конечно, только тогда, когда звезды имеют одинаковую яркость.

Теперь легко понять, почему меняется резкость полос при изменении расстояния между отверстиями на экране A. Если отверстия расположены очень близко друг к другу, то, как видно из формулы (4.4), фазы вторичных источников S_1 и S_2 будут практически совпадать друг с другом как для одной, так и для другой звезды. Интерференционная картина будет отчетливой. Если увеличивать расстояние между отверстиями, то интерференционные картины от разных звезд будут смещаться друг относительно друга, и при некотором расстоянии d_0 светлые полосы одной картины совпадут с темными полосами другой — интерференционная картина исчезнет. Пусть в некоторой точке P (рис. 4.1, *в*) находится светлая полоса одной интерференционной картины и темная — другой. Это означает, что световые колебания от одной звезды приходят в точку P в фазе, от другой — в противофазе. Поскольку разность хода лучей от вторичных источников S_1 и S_2 до точки P одинакова для обеих интерференционных картин, нетрудно сообразить, что наложение светлой полосы на темную имеет место при выполнении условия

$$l_1 + l_2 = (2k+1)\frac{\lambda}{2}, \qquad (4.5)$$

где k — любое целое число. С помощью формулы (4.4) условие полного исчезновения интерференционной картины

(4.5) можно записать в виде

$$d_k(\alpha_1+\alpha_2)=d_k\theta=(2k+1)\frac{\lambda}{2}.$$

Итак, при увеличении расстояния d между отверстиями первое исчезновение интерференционной картины происходит при

$$d=d_0=\frac{\lambda}{2\theta}. \qquad (4.6)$$

При дальнейшем увеличении d интерференционные полосы появляются снова, затем снова исчезают и т. д.

Измерив расстояние между отверстиями интерферометра d_0, при котором впервые исчезает интерференционная картина, мы получаем возможность с помощью формулы (4.6) вычислить угловой размер двойной звезды. Как видно из этой формулы, чувствительность прибора тем больше, чем больше может быть сделано расстояние между щелями на объективе.

Рис. 4.2. Схема звездного интерферометра Майкельсона.

Оценим, какой минимальный угловой размер двойной звезды можно измерить с помощью интерферометра на базе шестиметрового телескопа: при $\lambda=5500$ Å

$$\theta_{\min}=\frac{\lambda}{2D}\approx 0{,}01''.$$

Небольшим изменением рассмотренной выше конструкции звездного интерферометра Майкельсон сумел добиться высокого углового разрешения даже с помощью телескопа со сравнительно небольшим диаметром объектива. Он предложил свет от двойной звезды направлять через щели в объектив не непосредственно, как на рис. 4.1, *в*, а после отражения от системы зеркал, действие которых понятно из рис. 4.2. Расположенные против щелей S_1 и S_2 зеркала M_1 и M_2 закреплены неподвижно, а зеркала M_3 и M_4 можно симметрично раздвигать. При этом сдвиг интерференционной картины от одной звезды относительно картины от

другой и, следовательно, разрешающая способность прибора определяются расстоянием между зеркалами M_3 и M_4, хотя расстояние между интерференционными полосами зависит от расстояния между щелями и не меняется.

Рассмотрим теперь, какой вид будет иметь интерференционная картина, если вместо двух некогерентных точечных источников света имеется один протяженный источник с угловым размером θ. В этом случае каждый точечный элемент, на которые можно разбить протяженный источник, создает свою интерференционную картину. Так как все эти источники некогерентны, то их интерференционные картины просто налагаются друг на друга. Выясним, как меняется резкость суммарной картины от протяженного источника в виде светящейся полоски при постепенном увеличении расстояния d между щелями интерферометра. Если это расстояние очень мало, то положения полос интерференционных картин от всех точечных элементов светящейся полоски практически совпадают и суммарная картина имеет максимальную резкость. По мере увеличения расстояния d резкость картины уменьшается, и при некотором значении d_0 полосы пропадают. Найти величину d_0 можно следующим образом. Мысленно разобьем равномерно светящуюся полоску на пары одинаковых элементов так, чтобы расстояние между элементами любой пары равнялось половине длины полоски. Угловое расстояние между элементами каждой пары, очевидно, равно $\theta/2$. Если положение светлых полос интерференционной картины одного элемента пары совпадает с положением темных полос картины, создаваемой вторым элементом этой пары, то экран оказывается равномерно освещенным, так как условия совпадения одинаковы для всех пар элементов.

Из этих рассуждений следует, что условие исчезновения интерференционных полос от протяженного источника с угловым размером θ дается той же формулой (4.6), что и от двух точечных источников, только в ней следует заменить θ на $\theta/2$:

$$d_0 = \frac{\lambda}{\theta}. \tag{4.7}$$

В случае источника в виде равномерно светящегося диска условие исчезновения интерференционной картины

будет отличаться от (4.7) только лишь числовым множителем, близким к единице.

Звездный интерферометр Майкельсона позволяет определять не только угловое расстояние между компонентами двойных звезд, но и угловые диаметры не слишком удаленных одиночных звезд. Первой звездой, у которой Майкельсону удалось измерить угловой диаметр, была Бетельгейзе, относящаяся к так называемым красным гигантам. Он оказался равным 0,047″. Зная расстояние до Бетельгейзе, измеренное по параллаксу, можно было вычислить линейный диаметр звезды, оказавшийся равным примерно $4 \cdot 10^8$ км, что превышает диаметр земной орбиты ($3 \cdot 10^8$ км).

Формула (4.7) определяет допустимые размеры источника света при проведении интерференционных опытов по схеме Юнга: угловой размер источника θ, видимый от щелей в экране, не должен превышать отношения длины волны λ к расстоянию между щелями d.

§ 5. Интерференция немонохроматического света. Время когерентности

В этом параграфе мы рассмотрим изменения в интерференционных явлениях, которые вызываются отказом от монохроматической идеализации и учетом спектрального состава излучения реальных источников света.

Начнем с простейшего случая точечного источника, излучающего две очень узкие, близкие друг к другу спектральные линии с частотами ω_1 и ω_2. Если бы излучение на каждой из частот представляло собой бесконечную синусоиду, то результирующее излучение представляло бы собой волну средней частоты с периодически меняющейся амплитудой. Но в действительности вместо бесконечных синусоид излучаются более или менее длинные цуги волн определенной длины, причем начальные фазы колебаний в последовательно идущих цугах произвольны и никак не связаны друг с другом. Обычно за время наблюдения проходит много таких цугов, и поэтому излучения на частотах ω_1 и ω_2 можно считать независимыми. Другими словами, можно считать, что вместо одного имеется два расположенных в одном месте точечных источника, независимо друг от друга излучающих волны с частотами ω_1 и ω_2. При выполнении интерференционных опытов с таким

источником света каждая из волн создает свою интерференционную картину, и эти картины просто налагаются друг на друга.

Если частоты ω_1 и ω_2 мало отличаются друг от друга, то интерференционные полосы в каждой картине имеют почти одинаковую ширину. В тех местах, где светлые полосы одной картины налагаются на светлые полосы другой, резкость суммарной картины наибольшая. Наоборот, там, где светлые полосы одной картины приходятся на темные полосы другой, резкость интерференционных полос уменьшается вплоть до их полного исчезновения.

Найдем распределение освещенности в интерференционной картине, получаемой от двух вторичных источников, если первичный источник излучает две близкие спектральные линии одинаковой интенсивности. Интерференционная картина для отдельной спектральной линии была рассмотрена в § 1. Зависимость освещенности от разности хода l от вторичных источников до точки наблюдения дается формулой (1.5):

$$E(l) = 4E_0 \cos^2 \frac{\omega l}{2c} = 2E_0 \left(1 + \cos \frac{\omega l}{c}\right). \qquad (5.1)$$

Здесь E_0 — равномерная освещенность, которую создавал бы только один вторичный источник.

В рассматриваемом случае каждая спектральная линия первичного источника дает интерференционную картину, распределение освещенности в которой описывается формулой (5.1) с соответствующим значением частоты ω_1 или ω_2. Поэтому полное распределение освещенности, получающееся в результате наложения двух интерференционных картин, имеет следующий вид:

$$E(l) = E_1(l) + E_2(l) = \\ = 2E_{01}\left(1 + \cos \frac{\omega_1 l}{c}\right) + 2E_{02}\left(1 + \cos \frac{\omega_2 l}{c}\right). \qquad (5.2)$$

Поскольку спектральные линии имеют одинаковую интенсивность, то $E_{01} = E_{02} = E_0$ и формулу (5.2) можно преобразовать к виду

$$E(l) = 4E_0 \left(1 + \cos \frac{\Delta\omega\, l}{2c} \cos \frac{\omega l}{c}\right), \qquad (5.3)$$

где $\omega = {}^1/_2(\omega_1+\omega_2)$ — средняя частота, а $\Delta\omega=\omega_2-\omega_1$ — разность частот спектральных линий. Если частоты ω_1 и ω_2 близки, так что $\Delta\omega \ll \omega$, то резкость интерференционных полос медленно меняется с изменением разности хода l и распределение освещенности в зависимости от l имеет вид,

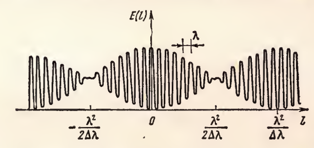


Рис. 5.1. Интерференционная картина в случае, когда источник света излучает две близкие спектральные линии.

показанный на рис. 5.1. Расстояние между соседними полосами определяется множителем $\cos\frac{\omega l}{c}$ и соответствует разности хода Δl, равной одной длине волны λ: $\frac{\omega}{c}\Delta l = 2\pi$, откуда $\Delta l = \frac{2\pi c}{\omega} = cT = \lambda$. Период изменения резкости полос определяется множителем $\cos\frac{\Delta\omega\, l}{2c}$ и соответствует разности хода Δl, равной произведению длины волны λ на отношение $\lambda/\Delta\lambda$. Действительно, как видно из рис. 5.1, период изменения резкости полос равен половине периода $\cos\frac{\Delta\omega}{2c}l$, поэтому $\frac{\Delta\omega\,\Delta l}{2c} = \pi$, откуда $\Delta l = \lambda^2/\Delta\lambda$.

Как можно наблюдать на опыте такую интерференционную картину с периодическим изменением резкости полос? Так как для этого необходима разность хода, равная очень большому числу длин волн, то наиболее удобно использовать интерферометр Майкельсона с подвижным зеркалом, схема которого приведена на рис. 1.5 (§ 1). Если плечи интерферометра почти равны друг другу, то наблюдаемые полосы соответствуют разностям хода, равным небольшому числу длин волн. При этом, как видно из рис. 5.1, полосы

имеют наибольшую резкость — освещенность на месте темных полос почти равна нулю. При перемещении зеркала разность хода l возрастает, а резкость интерференционных полос при этом постепенно убывает, так что при l порядка $\lambda^2/2\Delta\lambda$ полосы пропадают совсем. При дальнейшем перемещении зеркала полосы появляются снова, и при $l \sim \lambda^2/\Delta\lambda$ их резкость опять становится максимальной. Затем резкость снова убывает, и т. д.

Из изложенного ясно, что из наблюдения за изменением резкости интерференционных полос в зависимости от разности хода можно получить информацию о спектральном составе исследуемого света. Первые наблюдения такого рода были выполнены Физо в середине XIX века. В использованном им интерферометре наблюдались кольца Ньютона при освещении желтым светом натриевой лампы. Интерференционные полосы в данном случае имеют вид колец, так как разность хода волн, отразившихся от нижней поверхности линзы и верхней поверхности стеклянной пластинки (рис. 5.2), одинакова вдоль окружностей. Если линзу постепенно отводить от пластинки, то та же самая разность хода будет получаться на окружности меньшего радиуса, поэтому интерференционные кольца будут стягиваться к центру.

Рис. 5.2. К опыту Физо с кольцами Ньютона.

Физо нашел, что при контакте линзы с пластинкой кольца были резкими. При отодвигании линзы от пластинки резкость колец убывала, и при прохождении примерно 490-го кольца интерференционная картина исчезала. При дальнейшем увеличении расстояния кольца появлялись вновь и приобретали приблизительно первоначальную резкость при стягивании примерно 980-го кольца. Физо проследил периодическое изменение резкости полос в 52 периодах из 980 колец каждый! Отсюда он сделал правильный вывод о том, что натриевый свет состоит из двух спектральных линий почти равной интенсивности. Глядя на рис. 5.1, легко сообразить, что результаты опытов Физо дают для отношения $\lambda/\Delta\lambda$ у желтого дублета натрия значение, равное 980.

Рассмотренный пример света, состоящего из двух близких по частоте монохроматических волн, позволяет глубже проанализировать вопрос об использовавшейся в предыдущих параграфах монохроматической идеализации. Как известно, спектр испускания достаточно разреженных газов состоит из резких ярких линий, разделенных темными промежутками. Выделим свет одной из этих почти монохроматических линий и используем его в интерферометре Майкельсона. Мы увидим, что интерференционные полосы будут резкими, если длины путей обоих интерферирующих пучков примерно одинаковы. Если отодвигать одно из зеркал так, чтобы разность хода пучков увеличивалась, то резкость интерференционных полос будет постепенно уменьшаться, и в конце концов они исчезнут.

Такое исчезновение интерференционных полос легко объяснить, если считать, что свет излучается отдельными цугами, содержащими конечное число длин волн. Допустим для простоты, что все волновые цуги одинаковы. Каждый цуг, попадая в интерферометр, делится на два цуга равной длины. Если разность хода в плечах интерферометра больше этой длины, один из цугов минует точку наблюдения раньше, чем другой дойдет до нее, и интерференции наблюдаться не будет. Естественно ввести понятие длины когерентности как наибольшей разности хода интерферирующих лучей, при которой еще возможно наблюдение интерференционной картины. Длина когерентности характеризует степень отклонения рассматриваемого излучения от монохроматической идеализации и равна длине отдельных волновых цугов. Длину цуга волн можно характеризовать промежутком времени, в течение которого он проходит через точку наблюдения. Этот промежуток времени τ называется временем когерентности.

Однако исчезновение интерференционных полос при увеличении разности хода можно объяснить и на другом языке, рассматривая спектральный состав излучения. Строго монохроматической волне (бесконечной синусоиде) соответствует единственная частота, т. е. бесконечно узкая спектральная линия. Будем считать, что излучению, состоящему из волновых цугов конечной протяженности, соответствует спектральная линия некоторой конечной ширины. Другими словами, такое излучение можно рассматривать как совокупность отдельных монохроматических волн, частоты

которых сплошь заполняют некоторый интервал $\Delta\omega$, малый по сравнению со средней частотой ω. Каждая монохроматическая волна из этой совокупности создает в интерферометре свою интерференционную картину, и полное распределение освещенности определяется наложением этих картин.

При малых разностях хода интерферирующих лучей (порядка нескольких длин волн) положение интерференционных полос в картинах, создаваемых отдельными монохроматическими составляющими, будет практически совпадающим, и полосы суммарной картины будут отчетливыми. По мере увеличения разности хода отдельные интерференционные картины будут смещаться друг относительно друга, и в конце концов суммарная картина окажется полностью размытой. Оценить разность хода, при которой происходит исчезновение интерференционных полос, можно следующим образом. Мысленно разобьем весь спектральный интервал $\Delta\omega$, занимаемый рассматриваемым излучением, на пары монохроматических компонент, отстоящих друг от друга на $\Delta\omega/2$. Распределение освещенности от каждой пары дается формулой (5.3), в которой $\Delta\omega$ следует теперь заменить на $\Delta\omega/2$. Оно показано на рис. 5.1. Как видно из этого рисунка, полосы пропадают при такой разности хода l, когда аргумент первого косинуса в (5.3) становится равным $\pi/2$. Заменяя $\Delta\omega$ на $\Delta\omega/2$, получаем

$$\frac{\Delta\omega\, l}{4c}=\frac{\pi}{2}. \tag{5.4}$$

Условие исчезновения полос для всех пар монохроматических компонент одинаково. Поэтому при разности хода l, даваемой соотношением (5.4), происходит размытие полной интерференционной картины.

Теперь мы можем сопоставить две возможные интерпретации размывания интерференционных полос при достаточно большой разности хода — в рамках представлений о хаотичной последовательности ограниченных волновых цугов или представлений о суперпозиции монохроматических компонент, распределенных в некотором интервале частот. Так как в рамках представлений об отдельных цугах максимальная разность хода l равна длине цуга, то отношение l/c в соотношении (5.4) есть время когерентности τ. Переходя для удобства от циклической частоты ω к

частоте $\nu=\omega/2\pi$, переписываем (5.4) в виде

$$\tau\,\Delta\nu\approx1. \qquad (5.5)$$

Соотношение (5.5) следует рассматривать не как точное равенство, а только как оценку эффективного интервала частот $\Delta\nu$, границы которого в известной мере являются условными.

Мы видим, что чем больше длительность волновых цугов τ, тем у́же интервал частот $\Delta\nu$, в котором спектральные компоненты этого излучения имеют заметную величину. Иначе говоря, ширина спектральной линии излучения обратно пропорциональна времени когерентности.

Приведем оценки допустимых значений разности хода $l\sim\lambda^2/\Delta\lambda$ при наблюдении интерференции света с использованием разных источников. Для белого солнечного света или света, излучаемого раскаленными телами, интервал длин волн в спектре $\Delta\lambda$ одного порядка со средней длиной волны. Поэтому наблюдать интерференцию можно только при очень малых разностях хода, равных небольшому числу длин волн. Если воспользоваться излучением газоразрядной плазмы низкого давления, то при выделении какой-либо одной спектральной линии допустимая разность хода может быть значительно больше. Например, для красной линии кадмия с длиной волны $\lambda=6438$ Å, ширина которой $\Delta\lambda$ составляет всего лишь 0,013 Å, допустимая разность хода l превышает 500 000 длин волн, т. е. 30 см. А ширина линии излучения лазера может быть сделана настолько малой, что удается наблюдать интерференцию при разности хода в несколько километров!

§ 6. Физические принципы голографии

Голография — это способ записи и последующего восстановления световых волн, основанный на явлениях интерференции когерентных пучков света.

Разглядывая обычную фотографию, бессмысленно пытаться заглянуть за предметы, находящиеся на переднем плане. Это и естественно, так как фотография представляет собой плоское изображение объемной картины, полученное из определенной точки. В отличие от обычной фотографии, голография позволяет записать и восстановить не двумерное распределение освещенности в плоскости снимка, а

рассеянную предметом световую волну со всеми ее характеристиками — амплитудой, фазой, длиной волны. Само слово «голография» в буквальном переводе с греческого означает «полная запись». Восстановленные голограммой световые волны, попадая в глаз наблюдателя, создают полную иллюзию реальности наблюдаемых предметов — их объемность и возможность изменения ракурса при изменении точки зрения.

Рис. 6.1. Зонная пластинка Френеля.

Идеи, лежащие в основе голографической записи и восстановления зрительной информации, были высказаны английским физиком Габором в 1947 году. Так как для практической реализации голографии необходим свет с высокой степенью когерентности, то широкое распространение она получила после создания лазеров.

Чтобы понять принцип голографической записи и восстановления световых волн, рассмотрим действие так называемой зонной пластинки Френеля. Возьмем плоскую прозрачную пластинку, на которой нанесены концентрические окружности, радиусы которых r_k равны радиусам зон Френеля, видимым из некоторой точки P. Эти радиусы даются формулой (2.2) в параграфе «Дифракция света» (стр. 426). Таким образом, вся пластинка оказывается разбитой на зоны Френеля для некоторого значения длины волны λ. Теперь все нечетные (или, наоборот, четные) зоны Френеля должны быть сделаны непрозрачными. Это и есть зонная пластинка Френеля (рис. 6.1).

Предположим, что на зонную пластинку падает по нормали плоская монохроматическая волна длины λ. Тогда все прозрачные зоны Френеля в соответствии с принципом Гюйгенса можно рассматривать как источники когерентных вторичных волн. В точке P (рис. 6.2) эти вторичные волны будут, интерферируя, усиливать друг друга, так как разности хода между волнами, идущими от двух соседних прозрачных зон, равны длине волны λ. Точка P является, таким образом, тем фокусом, в котором сходятся волны, испытавшие дифракцию при прохождении через пластинку.

Но, кроме сходящейся в точке P сферической волны, в результате дифракции на зонной пластинке возникает также

расходящаяся сферическая волна, центр которой расположен в симметричной точке P' перед пластинкой (рис. 6.2): лучи *1*, *2*, ... будут восприниматься глазом как выходящие из одной точки P', так как разность хода между такими лучами, как ясно из рисунка, равна целому числу длин волн, что эквивалентно отсутствию разности хода вообще. Таким образом, точка P' представляет собой мнимый фокус

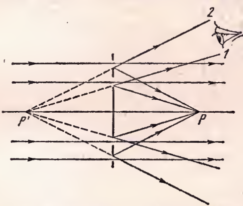


Рис. 6.2. К восстановлению изображения точечного источника с помощью голограммы.

расходящегося пучка лучей *1*, *2*, ..., возникающих в результате дифракции плоской волны на зонной пластинке.

Кроме сходящейся и расходящейся сферических волн, позади пластинки будет, разумеется, и плоская волна, так как у лучей, прошедших через прозрачные зоны пластинки без изменения направления, разность хода равна нулю.

Зонную пластинку Френеля можно получить фотографическим способом. Для этого достаточно светочувствительную пластинку поставить вместо экрана, на котором наблюдается интерференция плоской и сферической когерентных волн (рис. 6.3). Положение светлых и темных интерференционных полос на фотопластинке соответствует положению прозрачных и непрозрачных зон Френеля на зонной пластинке. В самом деле, в точках A_1 и A_2 колебания от плоской и сферической волн происходят в одинаковой фазе и дают соседние светлые полосы в интерференционной кар-

тине. Расстояния от точек A_1 и A_2 до центра S сферической волны различаются на одну длину волны λ. Но это как раз то условие, которому удовлетворяют соседние прозрачные кольца на зонной пластинке.

Представим себе, что на полученную таким образом фотографию падает только плоская волна, причем точно такая же, как и при фотографировании. Мы уже выяснили, что при этом за пластинкой будут наблюдаться три дифрагированные волны: плоская и две сферические. Лучи одной из сферических волн — сходящейся — пересекаются в точке P (рис. 6.2). Продолжения лучей другой сферической волны — расходящейся — пересекаются в точке P', положение которой совпадает с центром S сферической волны, использованной при фотографировании.

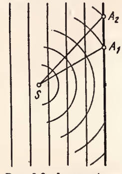


Рис. 6.3. Фотографическое получение зонной пластинки.

Эта расходящаяся сферическая волна и представляет наибольший интерес для голографии. Будем смотреть на пластинку из такого положения, чтобы в глаз попадала только расходящаяся волна (рис. 6.2). Тогда в глаз придет расходящийся пучок лучей, продолжения которых пересекаются в точке P', и мы увидим сквозь пластинку находящийся в P' точечный источник, хотя на самом деле никакого источника там нет!

Полученная описанным способом фотографическая пластинка с зонами Френеля и представляет собой голографическое изображение (голограмму) точечного источника монохроматического света: при дифракции плоской волны на этой голограмме происходит восстановление сферической волны точечного источника, использовавшегося при получении голограммы. Другими словами, расходящаяся сферическая волна, возникающая при дифракции плоской волны на голограмме, является точной копией волны, создававшейся точечным источником при записи голограммы.

При получении голограммы точки совершенно не обязательно, чтобы эта точка являлась источником света. Достаточно направить на нее свет, когерентный с плоской волной. Тогда голограмма образуется в результате интер-

ференции плоской волны, которую обычно называют опорной, и когерентной с ней сферической волны, рассеянной облучаемой точкой.

Результаты, полученные для одной точки, легко распространить на предметы любой формы, состоящие из большого числа точек, рассеивающих свет. На голограмме в этом случае получается сложный интерференционный узор, возникающий в результате интерференции опорной волны

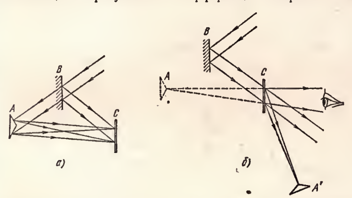


Рис. 6.4. Схемы получения (*а*) и восстановления (*б*) голограммы протяженного объекта.

и всех сферических вторичных волн, рассеянных отдельными точками предмета. При восстановлении в результате дифракции опорной волны на этом узоре возникают волны, расходящиеся от таких центров, где находились рассеивающие свет точки предмета при записи голограммы.

Схемы получения голограммы протяженного объекта и восстановления с ее помощью рассеянной объектом волны показаны на рис. 6.4, *а* и *б*. Плоская монохроматическая волна от лазера падает на полупрозрачную пластинку *B* и разделяется на две когерентные волны (рис. 6.4, *а*). Отраженная от *B* волна падает на фотопластинку *C* и играет роль опорной волны. Прошедшая сквозь *B* волна освещает объект *A*, каждая точка которого становится источником вторичных сферических волн. Рассеянные объектом волны также падают на фотопластинку *C*, где фиксируется результат их интерференции с опорной волной. Для получения

интерференционной картины на фотопластинке необходимо, чтобы время когерентности используемого света было большим: протяженность цуга волн должна быть больше разности хода между опорной волной и волной, рассеянной предметом. Именно поэтому необходимо использовать лазер. При восстановлении рассеянной волны (рис. 6.4, *б*) на голограмму падает та же опорная волна, которая использовалась при записи. Если расположить глаз (или фотоаппарат) позади голограммы, как указано на рисунке, то, воспринимая пучок расходящихся лучей от дифрагированной на голограмме опорной волны, наблюдатель увидит сквозь голограмму объект A в том месте, где он находился при записи голограммы. Разумеется, здесь, как и в рассмотренном выше случае точечного источника, кроме расходящихся волн будут также присутствовать прошедшая без отклонения волна и сходящиеся волны, которые дают действительное изображение объекта A'. Использование наклонного падения опорной волны приводит, как видно из рис. 6.4, *б*, к хорошему пространственному разделению всех трех волн, благодаря чему можно сквозь голограмму рассматривать мнимое изображение объекта A без помех со стороны других пучков. В отличие от обычной фотографии, здесь не используются ни линзы, ни другие устройства, формирующие изображения.

Система расходящихся волн, дающая мнимое изображение, неотличима от волн, исходивших от самого объекта. Благодаря этому голограмма полностью восстанавливает объемную структуру объекта и передает не только видимое пространственное расположение предметов, но и эффект параллакса, заключающийся в изменении видимого взаимного расположения предметов при перемещении точки наблюдения.

В отличие от обычной фотографии, голограмма содержит информацию об объекте в закодированной форме. Внешне голограмма ничем не напоминает этот объект. На глаз фотопластинка с голограммой представляется равномерно серой, и лишь в микроскоп можно увидеть сложный интерференционный узор. Еще одно отличие от обычной фотографии состоит в том, что для восстановления можно с равным успехом использовать и позитив, и негатив голограммы. Это легко понять, вспомнив, что в зонной пластинке, представляющей собой голограмму точечного источника, мож-

но сделать непрозрачными или четные, или нечетные зоны Френеля.

Любой участок голограммы содержит информацию обо всем объекте, в то время как различные участки обычной фотографии передают информацию только об отдельных его частях. Действительно, при записи голограммы на любую часть пластинки падают волны, рассеянные всеми частями объекта.

Объем информации, содержащейся на голограмме, значительно больше, чем на фотографии того же объекта. Если объект состоит из нескольких предметов, находящихся на разных расстояниях, то при фотографировании можно получить четкое изображение, строго говоря, только для одного из них. При восстановлении голограммы такого объекта все предметы будут наблюдаться вполне четкими при соответствующей аккомодации глаза.

ГЕОМЕТРИЧЕСКАЯ ОПТИКА

§ 7. Световые лучи. Принцип Ферма

Принцип Гюйгенса позволяет легко установить законы, описывающие поведение плоской волны на плоской границе раздела двух прозрачных сред. С его помощью по заданной волновой поверхности падающей волны можно построить волновые поверхности отраженной и преломленной волн. Волновые поверхности этих волн остаются плоскими. Для задания положения волновых поверхностей плоской волны удобно ввести лучи — линии, перпендикулярные к волновым поверхностям, которые одновременно указывают и направление распространения волны. Вытекающие из принципа Гюйгенса правила нахождения лучей для отраженной и преломленной волн — это хорошо известные законы отражения и преломления.

Плоская волна характеризуется тем свойством, что ее волновые поверхности представляют собой неограниченные плоскости, а направление ее распространения и амплитуда везде одинаковы. Часто электромагнитные волны, не являющиеся плоскими, можно приближенно рассматривать как плоские на небольшом участке пространства. Для этого необходимо, чтобы амплитуда и направление распро-

странения волны почти не менялись на протяжении расстояний порядка длины волны. Тогда также можно ввести понятие лучей, т. е. линий, касательная к которым в каждой точке совпадает с направлением распространения волны. Если при этом граница раздела двух сред, например поверхность линзы, может считаться приблизительно плоской на расстояниях порядка длины волны, то поведение лучей света на такой границе будет описываться теми же законами отражения и преломления.

Изучение законов распространения световых волн в этом случае составляет предмет геометрической оптики, поскольку в этом приближении оптические законы можно сформулировать на языке геометрии. Многие оптические явления, такие, как, например, прохождение света через оптические системы, формирующие изображение, можно рассматривать, исходя из представления о световых лучах, совершенно отвлекаясь от их волновой природы. Поэтому представления геометрической оптики справедливы лишь в той степени, в какой можно пренебречь явлениями дифракции световых волн. Дифракция сказывается тем слабее, чем меньше длина волны. Это значит, что геометрическая оптика соответствует предельному случаю малых длин волн: $\lambda \to 0$.

Физическую модель пучка световых лучей можно получить, если пропустить свет от источника пренебрежимо малого размера через небольшое отверстие в непрозрачном экране. Выходящий из отверстия свет заполняет некоторую область, и если длина волны пренебрежимо мала по сравнению с размерами отверстия, то на небольшом расстоянии от него можно говорить о пучке световых лучей с резкой границей.

Основные законы геометрической оптики — закон прямолинейного распространения света в однородной среде, законы отражения и преломления света на границе раздела двух сред — могут быть получены с помощью принципа Ферма. Согласно этому принципу действительный путь распространения монохроматического луча света есть путь, для прохождения которого свету требуется экстремальное (как правило, минимальное) время по сравнению с любым другим мыслимым путем между теми же точками. Такая формулировка принципа Ферма на самом деле не вполне верна. Согласно принципу Ферма оптический путь должен сравниваться не с любым другим, а с ближайшим. Иначе

принцип будет просто неверен, например, когда свет от источника может попасть в какую-либо точку и непосредственно, и после отражения от зеркала. Однако точная формулировка принципа Ферма выходит за рамки данной книги.

Поскольку скорость света в среде с показателем преломления n равна c/n, принцип Ферма можно сформулировать как требование минимальности оптической длины

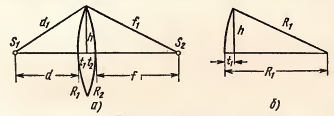


Рис. 7.1. К выводу формулы тонкой линзы.

луча при распространении света между двумя заданными точками. Под оптической длиной луча понимается произведение показателя преломления на длину пути луча. В неоднородной среде оптическая длина складывается из оптических длин на отдельных участках. Использование этого принципа позволяет рассмотреть некоторые задачи с несколько иной точки зрения, чем при непосредственном применении законов отражения и преломления. Например, при рассмотрении фокусирующей оптической системы вместо применения закона преломления можно просто потребовать равенства оптических длин всех лучей.

Получим с помощью принципа Ферма формулу тонкой линзы, не прибегая к закону преломления. Для определенности будем рассматривать двояковыпуклую линзу со сферическими преломляющими поверхностями, радиусы кривизны которых равны R_1 и R_2 (рис. 7.1).

Хорошо известно, что с помощью собирающей линзы можно получить действительное изображение точки. Пусть S_1 — предмет, S_2 — его изображение. Все лучи, исходящие из S_1 и прошедшие через линзу, собираются в одной точке S_2. Пусть S_1 лежит на главной оптической оси линзы, тогда изображение S_2 также лежит на оси. Что значит получить формулу линзы? Это значит установить связь между

расстояниями d от предмета до линзы и f от линзы до изображения и величинами, характеризующими данную линзу: радиусами кривизны ее поверхностей R_1 и R_2 и показателем преломления n. Из принципа Ферма следует, что оптические длины всех лучей, выходящих из источника и собирающихся в точке, являющейся его изображением, одинаковы. Рассмотрим два из этих лучей: один, идущий вдоль оптической оси, второй — через край линзы (рис. 7.1, a). Несмотря на то, что второй луч проходит большее расстояние, его путь в стекле короче, чем у первого, так что время распространения света от S_1 до S_2 для них одинаково. Выразим это математически. Обозначения величин всех отрезков указаны на рисунке. Приравняем оптические длины первого и второго лучей:

$$d+n(t_1+t_2)+f=d_1+f_1. \tag{7.1}$$

Выразим d_1 по теореме Пифагора:

$$d_1=\sqrt{(d+t_1)^2+h^2}=(d+t_1)\sqrt{1+\frac{h^2}{(d+t_1)^2}}.$$

Теперь воспользуемся приближенной формулой $\sqrt{1+x}\approx 1+x/2$, которая справедлива при $x\ll 1$ с точностью до членов порядка x^2. Считая h малым по сравнению с d, с точностью до членов порядка $(h/d)^4$ имеем

$$d_1\approx d+t_1+\frac{1}{2}\frac{h^2}{d+t_1}. \tag{7.2}$$

Аналогично для f_1 получаем

$$f_1\approx f+t_2+\frac{1}{2}\frac{h^2}{f+t_2}. \tag{7.3}$$

Подставляем выражения (7.2) и (7.3) в основное соотношение (7.1) и приводим подобные члены:

$$(n-1)(t_1+t_2)=\frac{h^2}{2}\left(\frac{1}{d+t_1}+\frac{1}{f+t_2}\right). \tag{7.4}$$

В этой формуле в случае тонкой линзы можно пренебречь величинами t_1 и t_2 в знаменателях правой части по сравнению с d и f; очевидно, что в левой части выражения (7.4) t_1+t_2 следует сохранить, ибо этот член стоит множителем.

С той же точностью, что и в формулах (7.2) и (7.3), t_1 и t_2 с помощью теоремы Пифагора можно представить в

виде (рис. 7.1, *б*)

$$t_1 = R_1 - \sqrt{R_1^2 - h^2} \approx \frac{h^2}{2R_1}, \quad t_2 \approx \frac{h^2}{2R_2}.$$

Теперь остается только подставить эти выражения в левую часть формулы (7.4) и сократить обе части равенства на $h^2/2$:

$$(n-1)\left(\frac{1}{R_1} + \frac{1}{R_2}\right) = \frac{1}{d} + \frac{1}{f}.$$

Это и есть искомая формула тонкой линзы. Вводя обозначение

$$\frac{1}{F} = (n-1)\left(\frac{1}{R_1} + \frac{1}{R_2}\right), \tag{7.5}$$

ее можно переписать в виде

$$\frac{1}{d} + \frac{1}{f} = \frac{1}{F}. \tag{7.6}$$

Нетрудно понять, что F есть фокусное расстояние линзы: если источник находится на бесконечности (т. е. на линзу падает параллельный пучок лучей), его изображение находится в фокусе. Полагая $d \to \infty$, получаем $f \to F$.

Полученное свойство фокусировки параллельного пучка монохроматических лучей является, как видно из проделанного вывода, приближенным и справедливо лишь для узкого пучка, т. е. для лучей, не слишком сильно отстоящих от оптической оси. Для широких пучков лучей имеет место сферическая аберрация, т. е. далекие от оптической оси лучи пересекают ее не в фокусе (рис. 7.2). В результате изображение бесконечно удаленного точечного источника, создаваемое широким пучком лучей, преломленных линзой, оказывается несколько размытым.

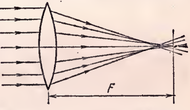


Рис. 7.2. Сферическая аберрация линзы.

Кроме сферической аберрации, линза как оптический прибор, формирующий изображение, обладает рядом других недостатков. Например, даже узкий параллельный

пучок монохроматических лучей, образующий некоторый угол с оптической осью линзы, после преломления не собирается в одну точку. При использовании немонохроматического света у линзы проявляется еще и хроматическая аберрация, связанная с тем, что показатель преломления n зависит от длины волны. В результате, как видно из формулы (7.5), узкий параллельный пучок лучей белого света пересекается после преломления в линзе не в одной точке: лучи каждого цвета имеют свой фокус. При конструировании оптических приборов удается в большей или меньшей степени устранить те или иные недостатки путем применения специально рассчитанных сложных многолинзовых систем. Однако одновременно устранить все недостатки невозможно. Поэтому приходится идти на компромисс и, рассчитывая оптические приборы, предназначенные для определенной цели, добиваться устранения одних недостатков и мириться с присутствием других. Например, объективы, предназначенные для наблюдения объектов малой яркости, должны пропускать возможно больше света, что вынуждает мириться с некоторыми аберрациями, неизбежными при использовании широких пучков света.

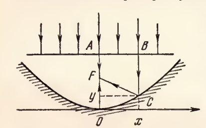


Рис. 7.3. Все параллельные лучи после отражения от параболического зеркала собираются в точке F.

Для объективов телескопов, где изучаемыми объектами являются звезды — точечные источники, расположенные вблизи оптической оси прибора, особенно важно устранить сферическую и хроматическую аберрацию для широких пучков, параллельных оптической оси. Устранить хроматическую аберрацию проще всего путем использования в оптической системе отражения вместо преломления. Так как лучи всех длин волн отражаются одинаково, то телескоп-рефлектор, в отличие от рефрактора, полностью лишен хроматической аберрации. Если при этом еще надлежащим образом выбрать форму поверхности отражающего зеркала, то можно полностью избавиться и от сферической аберрации для пучков, параллельных оптической оси. Для получения

точечного осевого изображения зеркало должно быть параболическим.

Покажем это. Пусть плоская волна, т. е. пучок лучей, параллельных оси y, падает на зеркальную поверхность, обладающую тем свойством, что после отражения все лучи собираются в одной точке F (рис. 7.3). Из симметрии ясно, что искомая поверхность зеркала представляет собой поверхность вращения вокруг оси y, поэтому достаточно рассмотреть сечение этой поверхности плоскостью xy, т. е. кривую $y=y(x)$. Рассмотрим центральный луч и луч, падающий на зеркало в произвольной точке C с координатами x и y. На основании принципа Ферма оптическая длина этих лучей от произвольной волновой поверхности AB до фокуса F должна быть одной и той же:

$$BC+CF=AO+OF. \qquad (7.7)$$

Из рис. 7.3 видно, что $AO=BC+y$, а $CF=\sqrt{x^2+(F-y)^2}$. Подставляя эти значения в (7.7), получим

$$\sqrt{x^2+(F-y)^2}=y+F.$$

Возводя обе части в квадрат и приводя подобные члены, найдем

$$y=\frac{1}{4F}x^2. \qquad (7.8)$$

Это уравнение параболы.

Параболические зеркала используются во всех крупнейших телескопах, в том числе и в самом большом 6-метровом телескопе. В этих телескопах устранены сферическая и хроматическая аберрации, однако параллельные пучки, идущие даже под небольшими углами к оптической оси, после отражения не пересекаются в одной точке и дают сильно искаженные внеосевые изображения. Поэтому пригодное для работы поле зрения оказывается очень небольшим, порядка нескольких десятков угловых минут.

Геометрическая оптика позволяет сравнительно просто рассмотреть прохождение света через оптическую систему. Однако использование законов геометрической оптики в некоторых случаях может привести к неверным результатам из-за проявления светом волновых свойств. Например, вблизи точки пересечения пучка лучей искривление волновой поверхности становится настолько существенным, что

ее уже нельзя считать плоской на расстояниях порядка длины волны. Вблизи таких точек условия применимости геометрической оптики заведомо не выполняются: световой поток нельзя собрать в одну точку, ибо это привело бы к бесконечно большой освещенности, чего на самом деле не бывает.

В какой мере волновые свойства света искажают предсказываемую геометрической оптикой картину, можно увидеть на примере простейшего оптического прибора — камеры-обскуры.

Устройство камеры-обскуры схематически показано на рис. 7.4. Она представляет собой ящик, в одной из стенок которого сделано малое отверстие. Действие камеры-обскуры, как и существование резких теней от непрозрачных предметов при малом источнике света,— это факты, указывающие на прямолинейное распространение света в однородной среде. Однако основной закон геометрической оптики — прямолинейное распространение света — строго справедлив лишь для широких, строго говоря, неограниченных световых пучков. Всякое ограничение ширины светового пучка, неизбежное в любом оптическом приборе, обязательно приводит к отступлениям от геометрической оптики и к проявлениям волновых свойств света. Выбор оптимального диаметра отверстия для получения на экране наиболее резкого изображения удаленных предметов — это поиск определенного компромисса между волновой и геометрической оптикой. Если бы свет действительно подчинялся законам геометрической оптики, то задача была бы тривиальной: чем меньше отверстие, тем резче изображение. В самом деле, удаленный предмет можно разбить на отдельные элементы и каждый элемент рассматривать как точечный источник. Отверстие в передней стенке камеры вырезает пучок лучей от источника, попадающих на экран. Пучок лучей от удаленной светящейся точки можно считать параллельным; поэтому размер пятна на экране, которое мы рассматриваем как изображение этой точки, определяется размером отверстия. При оценке размер пятна можно считать равным размеру отверстия. Но уменьшать отверстие

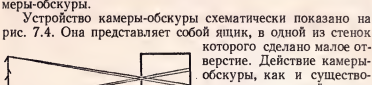


Рис. 7.4. Схема камеры-обскуры.

беспредельно нельзя не только потому, что при этом уменьшается световой поток и, следовательно, освещенность изображения, но и потому, что рано или поздно начнет сказываться волновая природа света. Дифракция света на отверстии приводит к размыванию изображения. Если уменьшать отверстие до размеров, сравнимых с длиной волны света, то изображение исчезает совсем и экран становится практически равномерно освещенным.

Оценим размер дифракционного пятна на экране, которое можно рассматривать как изображение удаленного точечного источника, в тех случаях, когда необходимо пользоваться волновой оптикой. Это можно сделать точно так же, как в § 4, где оценивались размеры дифракционного изображения звезды в телескопе. Согласно формуле (4.1) (стр. 443), для угла дифракции θ, т. е. направления на край центрального дифракционного пятна, имеем

$$\theta = \frac{\lambda}{d},$$

где d — диаметр отверстия камеры-обскуры. Этот угол определяет линейный размер a дифракционного пятна на экране камеры-обскуры. Если расстояние от отверстия до экрана равно L, то

$$a \approx 2L\theta = 2\frac{\lambda}{d}L.$$

Очевидно, что уменьшать размер отверстия следует только до тех пор, пока величина дифракционного пятна не сравняется с размером изображения, получающегося в приближении геометрической оптики. Дальнейшее уменьшение отверстия приведет только к размыванию изображения, т. е. к ухудшению резкости.

Итак, наилучшая резкость изображения достигается при равенстве диаметра отверстия и размера дифракционного пятна a:

$$d \approx 2\frac{\lambda}{d}L, \quad \text{откуда} \quad d \approx \sqrt{2\lambda L}.$$

При $L = 25$ см для видимого света ($\lambda \approx 5 \cdot 10^{-5}$ см) оптимальный размер отверстия равен 0,5 мм.

§ 8. Оптические приборы для визуальных наблюдений. Телескоп

Кажущаяся величина рассматриваемого предмета определяется величиной его изображения на сетчатке глаза. Размер изображения на сетчатке зависит от угла, под которым виден предмет. Определение угла зрения θ ясно из рис. 8.1. Угол зрения не может быть меньше некоторого минимального значения, примерно равного 1′, в противном случае глаз не может разрешить две точки, т. е. видеть их раздельно.

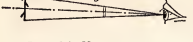


Рис. 8.1. Угол зрения.

Угол зрения можно увеличить, приближая глаз к предмету. Для нормального глаза имеет смысл приближать предмет не более чем до 25 см, т. е. до расстояния наилучшего зрения, наиболее удобного для рассматривания деталей предмета. При меньших расстояниях человек с нормальным зрением лишь с трудом аккомодирует свой глаз. Но если

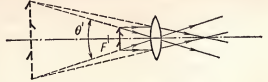


Рис. 8.2. Ход лучей в лупе.

перед глазом поместить собирательную линзу (лупу), то рассматриваемый предмет можно значительно приблизить к глазу и тем самым увеличить угол зрения. Отношение угла зрения при наблюдении предмета через оптический прибор к углу зрения при наблюдении невооруженным глазом на расстоянии наилучшего зрения называется увеличением прибора.

Ход лучей при рассматривании предмета через лупу показан на рис. 8.2. Предмет помещен перед линзой на расстоянии, немного меньшем фокусного. Лучи от любой точки предмета после преломления в линзе образуют пучок расходящихся лучей, продолжения которых пересекаются в одной точке, создавая мнимое изображение. Это изображение рассматривается глазом, помещаемым непосредст-

вению за лупой. При небольшом перемещении предмета вблизи фокуса положение мнимого изображения меняется значительно, и при совмещении предмета с фокусом его мнимое изображение вообще удаляется на бесконечность. Однако угловой размер θ' изображения, как можно видеть из рис. 8.2, при этом почти не меняется. Поэтому положение предмета практически не влияет на увеличение лупы, а сказывается только на аккомодации глаза при рассматривании мнимого изображения. Легко видеть, что увеличение лупы равно отношению расстояния наилучшего зрения d_0 к фокусному расстоянию F:

$$\Gamma = \frac{d_0}{F}. \tag{8.1}$$

Но иногда приблизиться к предмету невозможно. Так обстоит дело, например, при наблюдении небесных тел. Тогда

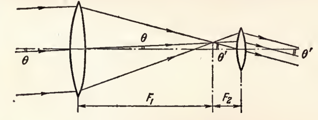


Рис. 8.3. Ход лучей в телескопе.

с помощью большой линзы, называемой объективом, получают действительное изображение удаленного тела. Это изображение значительно меньше, чем сам предмет, но зато к нему можно приблизить глаз и тем самым увеличить угол зрения. Так получается телескоп с одной линзой. Если же это изображение рассматривать в лупу (называемую окуляром), то можно еще больше приблизить глаз к действительному изображению удаленного предмета и тем самым еще больше увеличить угол зрения. Ход лучей в простейшем двухлинзовом телескопе показан на рис. 8.3. От каждой точки удаленного предмета в объектив приходит практически параллельный пучок лучей, который дает изображение этой точки в фокальной плоскости объектива. Чтобы при наблюдении не напрягать глаз, фокальную плоскость лупы (окуляра) обычно совмещают с фокальной плоскостью

объектива. Тогда падающий на объектив параллельный пучок лучей выходит из окуляра также параллельным. Пусть невооруженным глазом предмет виден под углом θ. Отношение угла θ', под которым предмет виден в телескоп, к углу θ называется увеличением телескопа. Из рис. 8.3 видно, что это увеличение равно отношению фокусных расстояний объектива F_1 и окуляра F_2:

$$\Gamma = \frac{\theta'}{\theta} = \frac{F_1}{F_2}. \tag{8.2}$$

Для получения большого увеличения нужен длиннофокусный объектив и короткофокусный окуляр. Уменьшая фокусное расстояние окуляра, можно получить с данным объективом большее увеличение. Однако не всегда следует стремиться только к получению большого увеличения. Это целесообразно лишь тогда, когда мы рассматриваем яркий объект,

Рис. 8.4. К определению светового потока, попадающего в глаз наблюдателя.

излучающий много света. В случае слабо освещенных объектов требования иные. Предположим, что мы рассматриваем не точечные тела, такие, как звезды, а протяженные, например поверхность планеты. Нужно, чтобы освещенность изображения, получаемого на сетчатке глаза, была бы по возможности большей. Легко убедиться, что освещенность изображения протяженного объекта при наблюдении в телескоп не может быть больше, чем при наблюдении невооруженным глазом. В самом деле, если увеличение телескопа равно Γ, то площадь изображения на сетчатке глаза в Γ^2 раз больше, чем при наблюдении без телескопа. Какой максимальный световой поток может попасть в глаз при данном увеличении? Диаметр попадающего в глаз параллельного пучка лучей не может быть больше диаметра зрачка глаза d. Поэтому, как видно из рис. 8.4, пучок попадающих в глаз лучей перед телескопом не может иметь диаметр, больший

чем $D = \Gamma d$. Так как световой поток пропорционален квадрату диаметра пучка, то при наблюдении в телескоп световой поток может вырасти не больше чем в Γ^2 раз по сравнению с наблюдением невооруженным глазом. Итак, и площадь изображения на сетчатке глаза, и падающий на эту площадь световой поток вырастают в Γ^2 раз, и если можно

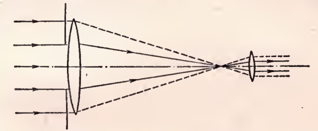


Рис. 8.5. К определению нормального увеличения.

пренебречь потерями света при отражении и поглощении в линзах, то освещенность изображения не меняется.

Из приведенных рассуждений ясно, что для получения заданного увеличения Г следует использовать объектив определенного диаметра, превосходящего диаметр зрачка глаза в Г раз. Если взять объектив большего диаметра, то часть собираемого им светового потока, как видно из рис. 8.5, просто не будет попадать в глаз. Если же взять объектив меньшего диаметра, то при прежнем увеличении уменьшится попадающий в глаз световой поток и освещенность изображения станет меньше. Это же можно сформулировать и иначе: для объектива заданного диаметра независимо от его фокусного расстояния существует определенное оптимальное увеличение, которое называется нормальным. Это есть то наибольшее увеличение, при котором получается изображение максимально возможной освещенности.

Таким образом, телескоп и глаз наблюдателя образуют единую систему, все элементы которой должны быть согласованы друг с другом. Это всегда учитывается при конструировании оптических приборов. Например, если мы хотим иметь полевой бинокль с десятикратным увеличением, то диаметр линз объектива должен быть в 10 раз больше диаметра зрачка глаза. Если принять средний диаметр зрачка равным 5 мм, то объектив должен быть диаметром 5 см.

Диаметр зрачка глаза не является постояиной величиной, он меняется от 6—8 мм в полной темноте до 2 мм при ярком диевном освещеиии. Поэтому при работе с телескопом, имеющим определеииый диаметр объектива, иапример 200 мм, нужно всегда учитывать обстаиовку, определяющую величииу зрачка глаза. Если наблюдается слабый объект в темиую иочь, когда зрачок имеет диаметр ие меиее 6 мм, целесообразно выбрать окуляр так, чтобы увеличеиие телескопа равиялось $\Gamma=200/6=33{,}4$. Но при иаблюдеиии

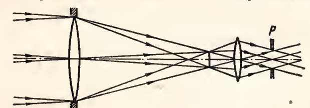


Рис. 8.6. При иаблюдеиии в телескоп глаз следует располагать вблизи изображения оправы объектива P.

днем, когда диаметр зрачка около 2 мм, целесообразио повысить увеличеиие втрое. Если фокусиое расстояиие F_1 иашего объектива равно 3 м, то в первом случае требуется окуляр с фокусиым расстояиием $F_2=300/33{,}4=9$ см, а во втором — 3 см.

При иаблюдеиии в телескоп протяжеииых объектов следует стремиться к тому, чтобы весь свет от объекта, входящий в объектив под разными углами, попадал бы в зрачок глаза. Для этого глаз следует располагать иа определеиием расстояиии от окуляра. В самом деле, окуляр, как собирательиая лииза, дает действительиое изображеиие оправы объектива телескопа. Так как в телескопе всегда $F_1 \gg F_2$, то это изображеиие P иаходится почти в фокальиой плоскости окуляра (рис. 8.6). Очевидио, что лучи, попадающие в объектив под разиыми углами, пройдут внутри этого изображеиия. Если условие согласования телескопа и глаза выполиено, то достаточио поместить зрачок глаза в то место, где иаходится изображение P оправы, чтобы все лучи попадали в глаз.

Астроиомические телескопы дают перевернутое изображеиие. Земные зрительиые трубы в основном подобиы астрономическим телескопам, за исключением того, что изобра-

жение у них должно быть прямым. Для переворачивания изображения можно воспользоваться либо призмами, как в полевом бинокле, либо дополнительными линзами.

Из-за волновой природы света изображение удаленной точки в фокальной плоскости объектива телескопа, как уже было показано, имеет вид дифракционного пятна. Изображения двух точек в фокальной плоскости объектива могут быть разрешены, если угловое расстояние между ними, как следует из формулы (4.3), не меньше значения $\theta \approx \lambda/D$. Каким следует выбрать увеличение телескопа, чтобы полностью использовать разрешающую способность его объектива? Пусть угловое расстояние между двумя удаленными точками как раз равно предельному значению λ/D, которое еще может разрешить объектив телескопа. В телескоп с увеличением Γ эти точки будут видны под углом $\alpha = \Gamma\lambda/D$. Чтобы эти точки воспринимались глазом как раздельные, этот угол не должен быть меньше угла $\beta \approx \lambda/d$, который способен разрешить глаз. Поэтому $\Gamma\lambda/D \geqslant \lambda/d$, откуда

$$\Gamma \geqslant \frac{D}{d}. \tag{8.3}$$

Знак равенства в этом выражении соответствует нормальному увеличению, при котором наиболее эффективно используется световой поток, попадающий в объектив телескопа. При увеличениях, меньших нормального, как мы видели, используется только часть объектива, что приводит к уменьшению разрешающей способности. Использование увеличений, бо́льших нормального, нецелесообразно, так как при этом разрешающая способность всей системы, определяемая пределом разрешения объектива λ/D, не увеличивается, а освещенность изображения на сетчатке глаза, как было показано выше, уменьшается.

Угловые размеры почти всех звезд много меньше разрешаемых угловых размеров даже самых больших телескопов. Поэтому изображение звезды в фокальной плоскости объектива телескопа неотличимо от изображения точечного источника света и представляет собой дифракционный кружок. Однако диаметр этого кружка настолько мал, что при использовании нормального увеличения он, как и сама звезда, для глаза неотличим от точечного источника света: размер дифракционного пятна на сетчатке глаза не зависит

от того, наблюдается ли звезда в телескоп или непосредственно. Если телескоп не отличает звезду от точечного источника, то в чем же его преимущество при наблюдении звезд по сравнению с невооруженным глазом? Дело в том, что в телескоп можно увидеть очень слабые звезды, вообще невидимые невооруженным глазом.

Так как размер дифракционного изображения звезды на сетчатке глаза не меняется при использовании телескопа, то освещенность этого изображения пропорциональна попадающему в глаз световому потоку. Но этот поток при использовании телескопа во столько раз больше светового потока, проходящего через зрачок невооруженного глаза, во сколько раз площадь отверстия объектива больше площади зрачка глаза.

ТЕОРИЯ ОТНОСИТЕЛЬНОСТИ

§ 9. Постулаты теории относительности. Принцип относительности. Максимальная скорость распространения взаимодействий

Проделав какой-нибудь эксперимент и повторив его при точно таких же условиях в другом месте и в другое время, мы получим тот же самый результат. Этот очевидный факт — воспроизводимость лабораторных опытов — находит свое выражение в независимости физических законов от таких обстоятельств, как положение в пространстве и выбор момента времени. Независимость явлений в замкнутой системе от места и момента времени является следствием однородности пространства и времени.

Опыт показывает, что, наряду с такой независимостью, существует определенная независимость физических явлений от состояния движения, которая заключается в равноправии всех инерциальных систем отсчета. Равномерное и прямолинейное движение замкнутой системы как целого не влияет на ход процессов, происходящих внутри системы. Утверждение об эквивалентности всех инерциальных систем отсчета составляет содержание принципа относительности. Этот принцип, впервые высказанный Галилеем для механических явлений, подтверждается всей совокупностью наших знаний о природе.

Миогие физические законы формулируются при помощи уравнений. Вид этих уравнеиий не зависит от иачального состояния системы. Таковы, в частности, уравнения механики, которые математически выражают второй закон Ньютона. Согласио принципу относительности математическая форма таких законов должна быть одинакова во всех инерциальных системах отсчета. Другими словами, уравнения движения должны быть инвариантны относительно перехода от одной инерциальной системы отсчета к другой.

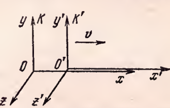


Рис. 9.1. Система отсчета K' движется относительно K со скоростью v вдоль оси x.

Рассмотрим описание некоторого явления в двух инерциальных системах отсчета K и K'. Система K' движется относительно K с постоянной скоростью $\boldsymbol{v}$. Условимся направление одноименных осей в K и K' выбирать одинаковым, а оси x и x' направим вдоль вектора $\boldsymbol{v}$ (рис. 9.1). Пусть иачало отсчета времени $t=0$ выбраио в тот момент, когда точки O и O' совпадали. Положение некоторой материальиой точки определяется координатами и временем x, y, z, t в системе K и кооординатами и временем x', y', z', t' в другой системе K'. Совокупность трех пространственных координат и времеии будем называть событием. Таким образом, событие, происходящее с иекоторой материальной частицей, определяется местом, где оно произошло, и времеием, когда оно произошло.

Как связаны между собой координаты и время иекоторого события, если его рассматривать в системах K и K'? В нерелятивистской физике принималось как очевидиый факт существование единого мирового времени t, одинакового во всех системах отсчета: $t=t'$. В действительности возможность измерять время во всех системах отсчета по одним и тем же часам связана с предположением о существовании сигналов, распространяющихся с бесконечно большой скоростью. Таким образом, согласно классическим представлениям, если два события происходят одновременно в некоторой системе отсчета, то они являются одновременными и в любой другой системе. Точно так же промежуток времеии между двумя событиями, в силу абсолютиого

характера времени, должен быть одинаковым во всех системах отсчета. Предполагалось также, а вернее, считалось очевидным, что длина твердого стержня или вообще расстояние между двумя точками, измеренное в некоторый момент времени, одинаковы во всех системах отсчета.

Из этих предположений однозначно вытекает общий вид преобразования, связывающего координаты и время некоторого события x, y, z, t в системе K с координатами и временем этого же события x', y', z', t' в системе K'. В самом деле, сравнивая координаты одной и той же частицы в системах отсчета K и K', немедленно получаем

$$x = x' + vt, \quad y = y', \quad z = z', \quad t = t'. \tag{9.1}$$

Эти формулы носят название преобразований Галилея.

Из преобразований Галилея можно сразу получить классический закон преобразования скорости частицы при переходе от одной системы отсчета к другой. Пусть $\boldsymbol{u} = d\boldsymbol{r}/dt$ — скорость некоторой частицы в K, а $\boldsymbol{u}' = d\boldsymbol{r}'/dt'$ — скорость той же частицы в K'. Поскольку $dt = dt'$, из преобразований Галилея получаем

$$u_x = u'_x + v, \quad u_y = u'_y, \quad u_z = u'_z. \tag{9.2}$$

Таким образом, преобразование скорости частицы при переходе от K к K' сводится просто к векторному сложению относительной и переносной скоростей, т. е. к сложению векторов $\boldsymbol{u}'$ и $\boldsymbol{v}$.

Уравнение движения классической механики $m\boldsymbol{a} = \boldsymbol{F}$ не меняет своего вида при переходе от одной инерциальной системы к другой, т. е. оно инвариантно относительно преобразований Галилея. Другими словами, преобразования Галилея удовлетворяют принципу относительности в отношении законов механики.

А как обстоит дело в электродинамике? Что говорит опыт о распространении принципа относительности на электромагнитные явления? Протекают ли электромагнитные и оптические процессы — взаимодействие зарядов и токов, распространение света — одинаково во всех инерциальных системах отсчета или равномерное прямолинейное движение лаборатории, не оказывая влияния на механические явления, сказывается на электромагнитных явлениях? Вся совокупность экспериментальных данных говорит о том, что принцип относительности распространя-

ется на все явления: как механические, так и электромагнитные и оптические процессы протекают одинаково во всех инерциальных системах отсчета. Исторически наиболее важные опыты, подтверждающие универсальный характер принципа относительности,— это электродинамический опыт Троутона и Нобля с заряженным конденсатором, подвешенным на упругой нити, и оптический опыт Майкельсона и Морли с интерферометром специальной конструкции. В этих опытах, поставленных специально для обнаружения влияния движения связанной с Землей лаборатории на взаимодействие зарядов и распространение света, был получен отрицательный результат: никакого влияния обнаружено не было. Однако уравнения электродинамики при переходе от одной инерциальной системы к другой, в отличие от уравнений динамики Ньютона, не являются инвариантными относительно преобразований Галилея. Простые соображения показывают, что преобразования Галилея не удовлетворяют принципу относительности в отношении законов электродинамики и оптики. В самом деле, согласно уравнениям Максвелла скорость распространения электромагнитных волн, в частности света, в вакууме одинакова по всем направлениям и равна $c=3\cdot10^{10}$ см/с. Но, с другой стороны, в соответствии с классическим законом преобразования скорости, вытекающим из преобразований Галилея, скорость света может быть по всем направлениям равна c только в одной инерциальной системе отсчета. Например, если скорость света равна c в системе K, то в K' свет должен распространяться в положительном направлении оси x' со скоростью $c-v$, а в отрицательном — со скоростью $c+v$. Отсюда можно сделать вывод, что уравнения электродинамики не инвариантны относительно преобразований Галилея.

Таким образом, между электродинамикой и классической механикой имеют место определенные противоречия. Опытные данные свидетельствуют о том, что принцип относительности распространяется на все явления, как механические, так и электродинамические и оптические. В то же время преобразования Галилея удовлетворяют принципу относительности в отношении законов механики и не удовлетворяют в отношении законов электродинамики и оптики.

На рубеже XIX и XX веков физика переживала глубокий кризис, единственно правильный выход из которого был

найден в 1905 году Эйиштейном ценой отказа от классических представлеиий о простраистве и времеии и от осиоваиных иа них преобразоваиий Галилея. Отказ от преобразоваиий Галилея и введеиие вместо иих новых преобразований (преобразоваиий Лореица), оставляющих иивариантиыми при переходе от одиой системы отсчета к другой уравнеиия электродииамики, а ие уравнения механики, требует пересмотра и уточиения закоиов классической механики. Решающим шагом иа этом пути оказался критический подход к используемому в классической физике поиятию абсолютиого времеии. Классические представлеиия, почерпиутые из повседиевного опыта и кажущиеся наглядиыми и очевидиыми, в действительиости оказались иесостоятельными. Миогие поиятия и величииы, которые в нерелятивистской физике считались абсолютиыми, т. е, ие зависящими от системы отсчета, теория отиосительности перевела в раиг отиосительиых. Например, считавшееся абсолютиым поиятие одиовремеииости двух событий в действительиости является отиосительиым: два удалеииых события, происходящие одновремеиио в некоторой системе отсчета, ие являются одновремеииыми в другой системе, движущейся отиосительио первой. Промежуток времеии между событиями, расстояиие между точками в простраистве — эти величииы также являются отиосительиыми.

Все физические явлеиия происходят в пространстве и во времеии, поэтому иеудивительно, что виесениое теорией отиосительности уточиение иекоторых основных поиятий, в особеииости воззреиий иа простраиствеиные и временные измереиия, затронуло в конечиом счете всю физику.

Теория отиосительиости основана иа двух принципах, или постулатах:

1) прииципˌ отиосительности;

2) прииципˌ существоваиия предельной скорости распространеиия взаимодействий.

Эти прииципы содержат настолько сильиые и общие утверждеиия, что едва ли возможио говорить о какихлибо «решающих» опытах, доказывающих их справедливость. Убеждение в справедливости этих принципов зиждется иа бесчислеииых опытиых проверках следствий теории относительиости, которая осиоваиа иа этих прииципах. Сюда относится вся совокупиость экспериментальиых даи-

ных, полученных при изучении движения быстрых частиц в приборах и ускорителях, атомных и ядерных процессов и т. п.

Принцип относительности, как уже отмечалось, есть утверждение об эквивалентности всех инерциальных систем отсчета. Равноправие всех инерциальных систем отсчета распространяется на все явления, на всю физику. Распространение принципа относительности на электромагнитные и оптические явления приводит к выводу о том, что скорость света (электромагнитных волн) в пустоте во всех инерциальных системах отсчета одинакова. Отсюда сразу видна необходимость пересмотра классических представлений о пространстве и времени, так как основанный на них классический закон преобразования скорости находится в противоречии с неизменностью скорости света.

Второй постулат утверждает, что любые взаимодействия между телами распространяются в пустоте с универсальной конечной скоростью, не зависящей от движения тел и равной скорости света в вакууме $c=3\cdot10^{10}$ см/с. Эта скорость определяет тот промежуток времени, после которого до тела может дойти первый сигнал, дающий знать об изменении, происшедшем с другим телом. Значение этого второго постулата связано с тем, что в определении понятий, относящихся к пространству и времени, фундаментальную роль играет передача сигналов с предельной скоростью. Передача сигналов в принципе возможна не только при помощи электромагнитных волн (света), но и при помощи волн другой природы. Мыслимо, хотя практически и неосуществимо, использование гравитационных волн. Не исключено открытие каких-либо новых полей, способных передавать сигналы. Можно, наконец, представить себе передачу сигналов при помощи предельно быстрых частиц. Однако принцип существования универсальной предельной скорости распространения взаимодействий утверждает существование общего предела для скорости передачи каких-либо действий и сигналов и придает скорости света в вакууме универсальное значение, не связанное с физической природой взаимодействия, а отражающее некоторое объективное свойство пространства и времени. Очевидно, что второй постулат утверждает в то же время, что невозможно движение тел со скоростью, превышающей предельную универсальную скорость c.

Отметим, что второй постулат находится в противоречии с принятым в классической механике способом описания взаимодействия материальных частиц, включающим в себя предположение о мгновенности распространения взаимодействий. В самом деле, силы, действующие на каждую из частиц со стороны остальных, считаются в классической механике зависящими от положения частиц в этот же момент времени. Изменение положения какой-либо из частиц мгновенно отражается на остальных. Поэтому второй постулат неизбежно требует уточнения законов механики. Механика теории относительности в предельном случае, когда скорости движущихся тел малы по сравнению со скоростью света *c*, переходит в классическую механику, основанную на мгновенности распространения взаимодействий. Только большой величиной скорости распространения взаимодействий объясняется тот факт, что для макроскопических тел в большинстве случаев достаточно точной оказывается классическая механика. В большинстве случаев скорости, с которыми приходится иметь дело, очень малы по сравнению с *c*. Поэтому в то время, когда была создана теория относительности, ее экспериментальное подтверждение можно было найти лишь в исключительно тонких оптических и электродинамических опытах. В настоящее время в больших ускорителях заряженные частицы нередко разгоняются до скоростей, составляющих 99% и более от скорости света. Для расчета траекторий столь быстрых частиц пользоваться механикой Ньютона уже нельзя. В этом смысле можно сказать, что теория относительности в наши дни стала инженерной наукой.

§ 10. Релятивистская кинематика. Синхронизация часов. Измерение промежутков времени и расстояний. Относительность промежутков времени и расстояний

Постулаты теории относительности требуют внесения изменений в основные физические понятия, относящиеся к пространству и времени. Прежде всего необходим анализ основных измерительных операций, определяющих пространственно-временные соотношения между событиями. Главное, что внесла теория относительности в постановку вопроса об измерительных операциях, состоит в том, что

любое физическое понятие (например, одновременность событий) и любая измерительная процедура (например, измерение промежутков времени и расстояний) нуждаются в определении.

Измерение времени может быть, в принципе, произведено при помощи любого периодического процесса. Наибольшей точностью в настоящее время обладают часы, основанные на использовании собственных колебаний молекул аммиака (молекулярные часы) или атомов цезия (атомные часы). Основанное на колебаниях атомов измерение времени особенно удобно тем, что здесь природа предоставила нам, в силу тождественности атомов одного и того же изотопа, набор совершенно идентичных часов. Выяснять, идут ли выбранные в качестве эталона часы равномерно, — бессмысленно: это по определению так.

Что значит измерить промежуток времени между событиями? Это значит сравнить между собой показания выбранных в качестве эталона часов в моменты наступления этих событий. А для этого нужно установить одновременность рассматриваемого события с другим событием — прохождением стрелки часов через определенное деление. Таким образом, все наши суждения, в которых время играет какую-либо роль, всегда являются суждениями об одновременных событиях. Понятие одновременности событий, происходящих в одном и том же месте, «рядом», не нуждается в определении. Поэтому для измерения промежутка времени между двумя событиями, происходящими в одном месте, достаточно иметь в этом же месте часы. Но как быть с «удаленными» событиями, происходящими в разных местах? Для измерения промежутка времени между такими событиями нужно иметь в тех точках, где они происходят, синхронно идущие идентичные часы.

Но как узнать, что находящиеся в различных точках A и B часы идут синхронно, или, что то же самое, как узнать, что два определенных события в точках A и B происходят одновременно? Узнать это нельзя, нужно сначала дать определение, что такое одновременность для пространственно удаленных точек. Без такого определения невозможно сравнивать по времени события, происходящие в различных точках.

Эйнштейновское определение одновременности основано на независимости скорости сигнала (распространяющегося

с предельной скоростью) от направления. Пусть из точки A в момент времени t_1 по часам в A отправляется сигнал (рис. 10.1). Пусть время прихода сигнала в B и отражения назад на часах в B есть t'. Наконец, пусть отраженный сигнал приходит в A в момент t_2 по часам в A. Тогда по определению часы в A и в B идут синхронно, если $t' = \frac{1}{2}(t_1 + t_2)$.

Как уже отмечалось, в нерелятивистской физике принималось как нечто само собой разумеющееся существование единого мирового времени, не зависящего от системы отсчета, и потому неявно допускалось, что понятие одновременности в разных точках пространства не нуждается в определении, а любой способ синхронизации часов (путем световых сигналов или путем перевозки хронометров) должен дать одно и то же. Мы увидим, что на самом деле это не так. Если часы в B синхронизированы с часами в A с помощью световых сигналов и хронометр C, сверенный с часами в точке A, перевозится затем в точку B, то его показания в точке B не совпадут с показаниями находящихся там часов, а будут зависеть от скорости перевозки. Совпадение будет лишь при бесконечно малой скорости перевозки хронометра.

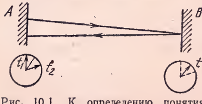


Рис. 10.1. К определению понятия одновременности событий.

Операцию измерения расстояний с точки зрения постулатов теории относительности наиболее разумно определить на основе «радиолокационного» способа: из некоторого пункта посылаются световые или радиосигналы, которые после отражения от наблюдаемого предмета возвращаются в точку отправления. При этом измеряется время прохождения сигнала туда и обратно по часам, связанным с радиолокатором. Расстояние до предмета получают, умножая одинаковую по всем направлениям скорость сигнала на половину времени прохождения: $l = \frac{1}{2} c (t_2 - t_1)$.

В принципиальном отношении этот способ важен потому, что в нем измерение расстояний сводится к измерению времени и отпадает необходимость в отдельном эталоне длины.

Радиолокационный способ измерения расстояний не является единственно возможным: можно было бы по определению в качестве измерительной процедуры принять способ непосредственного наложения жестких масштабов (линеек) или способ триангуляции. Однако эти способы менее удовлетворительны, так как существенно опираются

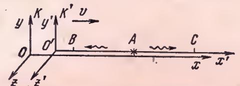


Рис. 10.2. Относительный характер одновременности событий.

на свойства твердых тел. Но абсолютно твердых тел в природе не существует, реальные физические тела могут рассматриваться как твердые и обладающие неизменными геометрическими размерами лишь приближенно.

До сих пор наши рассуждения относились к какой-либо одной инерциальной системе отсчета. Будем теперь рассматривать события, промежутки времени и расстояния с точки зрения разных систем отсчета.

В нерелятивистской физике время является абсолютным: для всех систем отсчета вводится одно и то же время. Это значит, что если два события происходят одновременно для какого-нибудь наблюдателя, то они являются одновременными и для любого другого: понятие одновременности является абсолютным, не зависящим от системы отсчета. Однако утверждение об абсолютном характере одновременности основано на предположении о существовании сигналов, распространяющихся мгновенно, с бесконечно большой скоростью. Покажем, что второй постулат теории относительности, утверждающий существование предельной *конечной* скорости сигналов, выражает относительный характер одновременности. Утверждение, что два пространственно удаленных события происходят одновременно, приобретает смысл только тогда, когда указано, к какой системе отсчета это утверждение относится.

Рассмотрим опять две инерциальные системы отсчета K и K', причем K' движется относительно K в положительном направлении оси x (рис. 10.2). Пусть из некоторой точки A

на оси x' отправляются сигналы во взаимно противоположных направлениях. Рассмотрим приход сигналов в точки B и C системы K', равноудаленные от A. Очевидно, что сигналы достигнут B и C одновременно по часам системы K'. Легко видеть, однако, что эти же два события — приход сигналов в B и C,— одновременные в K', отнюдь не будут одновременными для наблюдателя в системе K. В самом деле, согласно принципу относительности скорость сигналов в системе K также не зависит от направления, но точка B относительно K движется направо, навстречу посланному в нее сигналу, а точка C — по направлению от посланного в

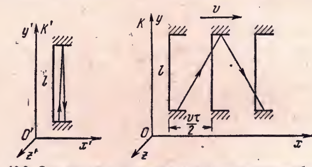


Рис. 10.3. Свет проходит от одного зеркала до другого и обратно за разное время в разных системах отсчета.

нее сигнала. Поэтому, с точки зрения наблюдателя в K, сигналу, распространяющемуся с конечной скоростью, приходится на пути в B преодолевать меньшее расстояние, чем на пути в C. Следовательно, в системе K сигнал в точку B придет раньше, чем в C, и, значит, понятие одновременности событий является относительным.

Покажем теперь, исходя из основных постулатов теории относительности, относительный характер промежутков времени между событиями. Пусть два события в некоторой системе отсчета, скажем в K', происходят в одной и той же точке и промежуток времени между ними равен τ_0 по часам системы K'. Этот промежуток времени называется собственным временем. Каким будет промежуток времени между этими же событиями, если его измерить по часам системы K, относительно которой K' движется со скоростью v? Для ответа на этот вопрос рассмотрим мысленный опыт со

«световыми часами», устроенными следующим образом (рис. 10.3). На концах стержня длиной l закреплены два параллельных зеркала. Между зеркалами движется короткий световой импульс. Пусть этот прибор неподвижен в системе K' и его стержень расположен перпендикулярно скорости K' относительно K. Рассмотрим один цикл таких часов, т. е. выход светового импульса от нижнего зеркала и его возвращение после отражения от верхнего зеркала, с точки зрения каждой из систем отсчета. В системе K' оба рассматриваемых события происходят в одной и той же точке и промежуток времени между ними (собственное время) равен $\tau_0=2l/c$. С точки зрения системы K часы находятся в движении и световой импульс движется между зеркалами зигзагообразно (рис. 10.3). Свет при этом проходит за один цикл больший путь, и, следовательно, промежуток времени τ между этими же событиями, измеряемый в системе отсчета K, больше, чем в K': $\tau>\tau_0$. В этих рассуждениях мы опираемся на то, что, согласно второму постулату, скорость света c одинакова в K и K'. Найдем связь τ и τ_0. Как видно из рис. 10.3, пройденный светом за один цикл путь равен $2\sqrt{l^2+(v\tau/2)^2}$, и для определения τ можно написать уравнение

$$c\tau=2\sqrt{l^2+\left(\frac{v\tau}{2}\right)^2},$$

откуда

$$\tau=\frac{2l}{c}\frac{1}{\sqrt{1-v^2/c^2}}.$$

Но, как мы видели выше, $2l/c$ равно промежутку времени τ_0 между этими событиями в K'. Поэтому

$$\tau=\frac{\tau_0}{\sqrt{1-v^2/c^2}}. \qquad (10.1)$$

Таким образом, величина промежутка времени между двумя событиями зависит от системы отсчета, т. е. является относительной. Так как при любой $v\neq0$ $\tau>\tau_0$, то собственное время меньше, чем промежуток времени между этими же событиями, измеренный в любой другой системе отсчета. Этот эффект называют релятивистским замедлением или «растяжением» времени. С точки зрения наблюдателя K идентичные по устройству движущиеся часы (т. е. часы в

K') идут медленнее, чем его собственные. Подчеркнем, что замедление времени является следствием инвариантности скорости света.

Рассмотренный релятивистский эффект замедления времени является взаимным, как того требует принцип относительности, т. е. постулат о равноправии инерциальных систем отсчета K и K': с точки зрения наблюдателя в системе K' медленнее идут часы, связанные с системой K.

Отметим, что в приведенных рассуждениях длину стержня l в направлении, перпендикулярном к относительной

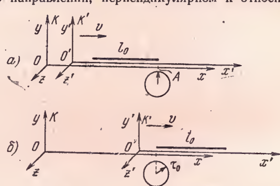


Рис. 10.4. Длина твердого стержня различна в разных системах отсчета.

скорости систем отсчета K и K', мы считали одинаковой в обеих системах отсчета. Если предположить, что это не так, то можно сразу прийти к противоречию с равноправием систем K и K'. В самом деле, рассмотрим следующий мысленный опыт. Расположим вдоль оси y' системы K' жесткий стержень, длина которого в этой системе равна l, и вдоль оси y системы K расположим точно такой же стержень, т. е. длина этого стержня равна l для наблюдателя в K. В некоторый момент эти стержни оказываются рядом, и представляется возможность сравнить их непосредственно — конец одного из стержней может сделать метку на другом стержне. Совпадет ли эта метка с концом стержня? Принцип относительности дает положительный ответ на этот вопрос: метка совпадет с концом стержня, т. е. длина стержня в направлении, перпендикулярном к относительной скорости

систем отсчета K и K', одинакова в обеих системах. Если бы совпадения не было, то один из стержней оказался бы длиннее другого с точки зрения обеих систем отсчета, что противоречит принципу относительности.

Покажем теперь, что длина твердого стержня, расположенного вдоль направления относительной скорости систем отсчета K и K' (рис. 10.4), будет различной в этих системах. Пусть стержень покоится в системе отсчета K'. Его длину, измеренную в этой системе отсчета, называют собственной длиной. Обозначим ее через l_0, а длину в системе K, относительно которой стержень движется со скоростью v, через l. Найдем связь между l и l_0. Для этого рассмотрим два события: *а*) прохождение начала стержня мимо точки A на оси x системы K и *б*) прохождение конца стержня мимо этой же точки. В системе K эти события происходят в одной точке, и промежуток времени между ними в системе K является собственным временем τ_0. Так как стержень движется относительно K со скоростью v, то можно написать: $l=v\tau_0$. Но с точки зрения наблюдателя в системе K' точка A движется вдоль неподвижного стержня налево с такой же по величине скоростью, поэтому $l_0=v\tau$, где τ есть промежуток времени между событиями (*а*) и (*б*), измеренный по часам в K'. Так как $\tau=\tau_0/\sqrt{1-v^2/c^2}$, то, комбинируя соотношения $l=v\tau_0$ и $l_0=v\tau$, находим

$$l=l_0\sqrt{1-v^2/c^2}. \qquad (10.2)$$

Мы приходим к выводу, что длина стержня зависит от системы отсчета, в которой она измеряется, т. е. является относительной. При любой $v\neq0$ $l<l_0$, т. е. длина стержня является наибольшей в той системе отсчета, в которой стержень покоится. Движущиеся относительно наблюдателя тела сокращаются в направлении своего движения. Этот релятивистский эффект носит название лоренцева сокращения. Лоренцево сокращение движущегося стержня отражает относительный характер расстояния между точками в теории относительности (т. е. зависимость расстояния от системы отсчета) и не связано с какими-либо процессами или явлениями в самом стержне, но тем не менее представляет собой вполне реальный эффект, столь же реальный, как, например, зависимость скорости тела от выбора системы отсчета.

В полном соответствии с принципом относительности эффект сокращения длины стержня является взаимным: если такой же стержень покоится в системе отсчета K, то его длина в этой системе отсчета равна l_0, а в системе K' длина будет меньше в соответствии с приведенной формулой.

Как видно из полученной формулы, эффект сокращения длины зависит от относительной скорости v систем отсчета и становится особенно заметным для скоростей, сравнимых со скоростью света. При $v \to c$ $l \to 0$. Зависимость лоренцева сокращения от скорости показана на рис. 10.5. При малых скоростях ($v \ll c$) $l \approx l_0$ и $\tau \approx \tau_0$, т. е. расстояние между точками и промежуток времени между событиями приобретают практически абсолютный смысл в полном соответствии с классическими представлениями о пространстве и времени, сформировавшимися на основе многовекового опыта наблюдений над сравнительно медленными движениями, происходящими со скоростями, малыми по сравнению со скоростью света.

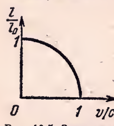


Рис. 10.5. Зависимость сокращения длины от относительной скорости.

§ 11. Преобразования Лоренца. Интервал. Релятивистский закон преобразования скорости

Полученные выше на основе постулатов теории относительности формулы (10.1) и (10.2), связывающие промежутки времени и расстояния между точками в разных системах отсчета, позволяют написать релятивистский закон преобразования координат и времени произвольного события при переходе от одной системы отсчета к другой. Этот закон должен заменить основанные на классических представлениях о пространстве и времени преобразования Галилея (9.1).

Рассмотрим, как и в § 9, описание некоторого события A в двух инерциальных системах отсчета K и K'. Пусть координаты и время этого события в системе K есть x, y, z и t, а в системе K' — x', y', z' и t' (рис. 11.1). Как и прежде, считаем, что при $t=0$ точки O и O' совпадают. Расстояния в направлении, перпендикулярном к относительной скорости $\boldsymbol{v}$ систем отсчета, как уже было показано, одина-

ковы в K и K', поэтому $y=y'$ и $z=z'$. Координата x есть собственная длина l_0 отрезка OB, неподвижного в K-системе. Длина l этого же отрезка в K'-системе, где измерение производится в момент времени t', есть $x'+vt'$. Учитывая соотношение (10.2) между собственной длиной некоторого отрезка l_0 и длиной l этого же отрезка в системе отсчета, относительно которой он движется со скоростью v:

$$l=l_0\sqrt{1-v^2/c^2},$$

мы можем написать

$$x'+vt'=x\sqrt{1-v^2/c^2},$$

откуда

$$x=\frac{x'+vt'}{\sqrt{1-v^2/c^2}}. \qquad (11.1)$$

Рис. 11.1. Координаты одного и того же события A в двух системах отсчета.

Но можно рассуждать и иначе. Координата x' есть собственная длина отрезка $O'B$, неподвижного в K'-системе. Длина этого же отрезка в K, измеряемая в момент времени t по часам K, равна $x-vt$. Снова учитывая соотношение (10.2) между длиной одного и того же отрезка в двух системах отсчета, можем написать

$$x'=\frac{x-vt}{\sqrt{1-v^2/c^2}}. \qquad (11.2)$$

Формулы (11.1) и (11.2) позволяют также найти связь между временем t и t' одного и того же события в обеих системах отсчета. Исключая из (11.1) и (11.2) сначала x', а затем x, найдем

$$t=\frac{t'+\frac{v}{c^2}x'}{\sqrt{1-v^2/c^2}}, \quad t'=\frac{t-\frac{v}{c^2}x}{\sqrt{1-v^2/c^2}}. \qquad (11.3)$$

Таким образом, релятивистские формулы преобразования координат некоторого события при переходе от одной инерциальной системы отсчета к другой имеют вид

$$x=\frac{x'+vt'}{\sqrt{1-v^2/c^2}}, \quad y=y', \quad z=z',$$
$$t=\frac{t'+\frac{v}{c^2}x'}{\sqrt{1-v^2/c^2}}. \qquad (11.4)$$

Эти формулы называют преобразованиями Лоренца. Они заменяют преобразования Галилея, справедливые лишь в предельном случае малых по сравнению со скоростью света относительных скоростей. При $v \ll c$ преобразования Лоренца (11.4) переходят в преобразования Галилея (9.1). Это означает, что теория относительности не отвергает полностью классические представления о пространстве и времени, а включает их в себя как предельный случай, справедливый для медленных движений. Теория относительности не отвергает классическую физику, а определяет границы ее применимости.

Преобразования Лоренца выражают относительный характер промежутков времени между событиями и расстояний между точками в пространстве. Однако наиболее характерной чертой теории относительности является не утверждение относительного характера пространства и времени, а установление абсолютных, не зависящих от выбора систем отсчета законов природы. Задача нахождения абсолютного выражения законов природы тесно связана с отысканием абсолютных, инвариантных величин. Одна из таких величин упоминается уже в основных постулатах — это максимальная скорость распространения взаимодействий, равная скорости света в вакууме c. Другой важной инвариантной величиной является пространственно-временной интервал между событиями, определяемый следующим соотношением:

$$S_{12} = \sqrt{c^2 t_{12}^2 - l_{12}^2}, \qquad (11.5)$$

где t_{12} — промежуток времени между событиями, а l_{12} — расстояние между точками, в которых происходят рассматриваемые события. В частности, если одно из событий происходит в начале координат $x_1 = y_1 = z_1 = 0$ в момент времени $t_1 = 0$, а второе — в точке x, y, z в момент t, то интервал между ними

$$S = \sqrt{c^2 t^2 - x^2 - y^2 - z^2}. \qquad (11.6)$$

Пусть, например, первое событие представляет собой вспышку света, происходящую в начале координат при $t = 0$, а второе — приход фронта этой световой волны в точку с координатами x, y, z в момент времени t. Тогда $x^2 + y^2 + z^2 = c^2 t^2$ и интервал для такой пары событий $S = 0$. Координаты и время второго события в другой системе отсчета

K' будут другими, но и для них в силу инвариантности скорости света будет выполняться такое же соотношение $x'^2+y'^2+z'^2=c^2t'^2$ и, следовательно, $S'=0$. Таким образом, если два события связаны между собой световым сигналом, то интервал между ними равен нулю во всех инерциальных системах отсчета. Этот результат является математическим выражением абсолютного характера скорости света.

Для любой другой пары событий, не связанных световым сигналом, интервал отличен от нуля, но величина его во всех инерциальных системах отсчета одинакова:

$$c^2t^2-x^2-y^2-z^2=c^2t'^2-x'^2-y'^2-z'^2.$$

В этом легко убедиться с помощью преобразований Лоренца (11.4), подставив в левую часть выражения для x, y, z и t через координаты и время этого же события x', y', z' и t' в другой системе отсчета. Инвариантность интервала означает, что утверждение «два события разделены интервалом S» имеет абсолютный характер, т. е. оно справедливо во всех инерциальных системах отсчета.

Понятие интервала между событиями является обобщением понятий промежутка времени и расстояния между точками. В зависимости от того, какая составляющая — временна́я или пространственная — преобладает в рассматриваемом интервале, возникает деление интервалов на времениподобные и пространственноподобные. Для времениподобного интервала $c^2t_{12}^2>l_{12}^2$, т. е. $S_{12}^2>0$. В этом случае всегда можно найти такую систему отсчета K', в которой рассматриваемые события происходят в одной точке, т. е. $l'_{12}=0$, и промежуток времени между ними в такой системе отсчета является собственным временем $t'_{12}=\tau_0$:

$$S_{12}^2=c^2t_{12}^2-l_{12}^2=c^2t_{12}'^2-l_{12}'^2=c^2\tau_0^2.$$

Таким образом, для событий, разделенных времениподобным интервалом, всегда существует такая система отсчета, в которой этот интервал (с точностью до постоянного множителя c) представляет собой просто промежуток времени τ_0 между этими событиями. Для этих событий понятия «раньше», «позже» имеют абсолютный характер. Очевидно, что между такими событиями может иметь место причинно-следственная связь.

Для событий, разделенных пространственноподобным интервалом, $c^2t_{12}^2<l_{12}^2$, т. е. $S_{12}^2<0$, а интервал является мнимым числом. В этом случае всегда можно найти такую систему отсчета K', в которой эти события происходят одновременно, т. е. $t'_{12}=0$:

$$c^2t_{12}^2-l_{12}^2=c^2t_{12}'^2-l_{12}'^2=-l_{12}'^2.$$

Абсолютная величина пространственноподобного интервала представляет собой пространственное расстояние между событиями в той системе отсчета, в которой эти события произошли одновременно. Понятия «одновременно», «раньше», «позже» для таких событий относительны: всегда можно указать такие системы отсчета, в которых первое событие происходит раньше второго, и такие, в которых второе происходит раньше первого. Ясно, что между событиями, для которых теряют абсолютный смысл понятия «раньше» и «позже», не может быть причинно-следственной связи. Впрочем, невозможность причинной связи между событиями, разделенными пространственноподобным интервалом, для которых $l_{12}>ct_{12}$, непосредственно видна из того, что никакой сигнал, никакое взаимодействие не может распространяться со скоростью, большей c.

Равный нулю интервал между событиями, связанными световым сигналом, называют светоподобным.

Подчеркнем еще раз, что разделение интервалов между событиями на времениподобные и пространственноподобные вследствие инвариантности интервала является абсолютным, т. е. не зависящим от системы отсчета.

Преобразования Лоренца для координат и времени события при переходе от одной инерциальной системы отсчета к другой позволяют сразу получить и закон преобразования скорости частицы. Пусть некоторая частица за малый промежуток времени Δt по часам системы отсчета K переместилась из точки с радиус-вектором $\boldsymbol{r}$ в точку с радиус-вектором $\boldsymbol{r}+\Delta\boldsymbol{r}$. Тогда по определению скоростью $\boldsymbol{u}$ этой частицы относительно системы K называется предел $\Delta\boldsymbol{r}/\Delta t$ при $\Delta t\to 0$. Скоростью $\boldsymbol{u}'$ этой же частицы в системе отсчета K' будет предел отношения $\Delta\boldsymbol{r}'/\Delta t'$, где $\Delta t'$ и $\Delta\boldsymbol{r}'$ — промежуток времени и изменение радиус-вектора в K' для той же пары событий, т. е. для конечного и начального состояний частицы. Применяя преобразования Лоренца (11.4) к конечному и начальному состояниям частицы и вычитая их

почленно, получим

$$\Delta x=\frac{\Delta x'+v\,\Delta t'}{\sqrt{1-v^2/c^2}},\quad \Delta y=\Delta y',\quad \Delta z=\Delta z',$$

$$\Delta t=\frac{\Delta t'+\frac{v}{c^2}\Delta x'}{\sqrt{1-v^2/c^2}}.$$

Разделив почленно первые три равенства на четвертое, находим

$$\frac{\Delta x}{\Delta t}=\frac{\frac{\Delta x'}{\Delta t'}+v}{1+\frac{v}{c^2}\frac{\Delta x'}{\Delta t'}},\quad \frac{\Delta y}{\Delta t}=\frac{\frac{\Delta y'}{\Delta t'}\sqrt{1-\frac{v^2}{c^2}}}{1+\frac{v}{c^2}\frac{\Delta x'}{\Delta t'}},$$

$$\frac{\Delta z}{\Delta t}=\frac{\frac{\Delta z'}{\Delta t'}\sqrt{1-\frac{v^2}{c^2}}}{1+\frac{v}{c^2}\frac{\Delta x'}{\Delta t'}}.$$

Переходя в этих формулах к пределу при $\Delta t\to 0$, $\Delta t'\to 0$, получим

$$u_x=\frac{u_x'+v}{1+\frac{v}{c^2}u_x'},\quad u_y=\frac{u_y'\sqrt{1-\frac{v^2}{c^2}}}{1+\frac{v}{c^2}u_x'},\quad u_z=\frac{u_z'\sqrt{1-\frac{v^2}{c^2}}}{1+\frac{v}{c^2}u_x'}.\quad (11.7)$$

Выражения (11.7) представляют собой закон преобразования скорости частицы при переходе от одной инерциальной системы отсчета к другой. Отметим, что поперечные к направлению относительной скорости систем отсчета компоненты скорости частицы u_y и u_z, в отличие от поперечных координат y и z, не остаются неизменными. Это связано с тем, что при переходе от одной системы отсчета к другой время преобразуется.

В предельном случае $v\ll c$ релятивистские формулы (11.7) переходят в формулы классической механики (9.2):

$$u_x=u_x'+v,\quad u_y=u_y',\quad u_z=u_z'.$$

Релятивистский закон преобразования скорости согласуется, разумеется, с исходным постулатом об инвариантности скорости света. Рассмотрим, например, в системе отсчета K' световой импульс, распространяющийся вдоль оси x'. Для такого импульса $u_x'=c$, $u_y'=u_z'=0$. Тогда,

согласно (11.7), для скорости этого же импульса в системе отсчета K получим

$$u_x = \frac{c+v}{1+\frac{v}{c}} = c, \quad u_y = 0, \quad u_z = 0$$

— импульс распространяется вдоль оси x со скоростью c.

В качестве примера применения релятивистского преобразования скорости рассмотрим точечный источник света,

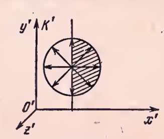


Рис. 11.2. Неподвижный источник света излучает равномерно по всем направлениям.

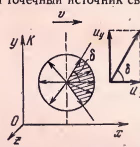


Рис. 11.3. Для наблюдателя, относительно которого источник движется, излучение не является изотропным.

покоящийся в системе K' и равномерно излучающий свет по всем направлениям (рис. 11.2). Рассмотрим те 50% светового потока, которые источник излучает в переднюю полусферу в системе отсчета K'. С точки зрения наблюдателя в системе K, относительно которой источник движется со скоростью v, излучение уже отнюдь не будет изотропным: эти 50% будут излучаться преимущественно вперед в конус с углом δ (рис. 11.3), причем $\cos\delta = v/c$. В самом деле, в системе K' луч света, ограничивающий рассматриваемый пучок, направлен вдоль оси y' и для него $u'_x = 0$, $u'_y = c$, $u'_z = 0$. Переходя в систему отсчета K, мы для этого же луча получим, согласно формулам (11.7),

$$u_x = v, \quad u_y = c\sqrt{1 - v^2/c^2}, \quad u_z = 0,$$

откуда

$$\cos\delta = \frac{u_x}{\sqrt{u_x^2 + u_y^2}} = \frac{v}{c}.$$

При скорости источника v, близкой к скорости света, 50% света сконцентрируется в узкий конус, направленный

вперед по движению, с осью, совпадающей с направлением движения источника.

Совершенно аналогично с помощью релятивистского закона преобразования скорости объясняется явление звездной аберрации. Движущийся наблюдатель обнаружит искажение картины звездного неба: для него Вселенная «сжата» в направлении его движения по сравнению с картиной, которую видит в том же направлении неподвижный относительно звезд наблюдатель. Если движущийся наблюдатель будет менять свою скорость, то он обнаружит, что звездное небо «переливается» вокруг него: направления, в которых он видит звезды, будут все время меняться, не образуя постоянных углов друг с другом. Именно в таком положении находится наблюдатель на Земле, обращающейся вокруг Солнца. Каждые полгода скорость Земли в ее годичном движении (30 км/с, т. е. $10^{-4}\,c$) меняет направление на противоположное. В явлении аберрации света звезд обнаруживается, конечно, не сама скорость Земли, а тот факт, что эта скорость изменяется: в разное время года положения звезд сдвинуты по-разному.

§ 12. Релятивистский импульс. Зависимость массы от скорости. Релятивистская энергия

Теория относительности требует пересмотра и уточнения законов механики. Как мы видели, уравнения классической динамики (второй закон Ньютона) удовлетворяют принципу относительности в отношении преобразований Галилея. Но ведь преобразования Галилея должны быть заменены преобразованиями Лоренца! Поэтому уравнения динамики следует изменить так, чтобы они оставались неизменными при переходе от одной инерциальной системы отсчета к другой согласно преобразованиям Лоренца. При малых скоростях ($v \ll c$) уравнения релятивистской динамики должны переходить в классические, ибо в этой области их справедливость подтверждается на опыте.

Прежде всего покажем, что релятивистское выражение для импульса частицы имеет вид $\boldsymbol{p}=m\boldsymbol{v}$, но теперь, в отличие от нерелятивистской механики, масса m зависит от скорости частицы следующим образом:

$$m=\frac{m_0}{\sqrt{1-v^2/c^2}}. \qquad (12.1)$$

Здесь m_0 — масса частицы в той системе отсчета, где она покоится, так называемая масса покоя частицы, совпадающая с массой частицы в нерелятивистской механике. Как мы увидим, зависимость массы от скорости частицы в теории относительности с неизбежностью следует из релятивистского эффекта замедления времени в движущейся системе отсчета.

Для выяснения зависимости массы частицы от скорости рассмотрим картину абсолютно упругого «скользящего»

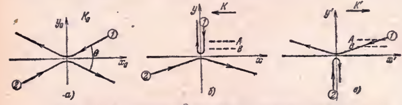


Рис. 12.1. К выводу зависимости массы тела от скорости.

столкновения двух одинаковых частиц. В системе центра масс это столкновение имеет вид, показанный на рис. 12.1, *а*: до столкновения частицы *1* и *2* движутся навстречу друг другу с одинаковыми по величине скоростями, после столкновения частицы разлетаются в противоположные стороны с такими же по величине скоростями, как и до столкновения. Другими словами, при столкновении происходит только поворот векторов скоростей каждой из частиц на один и тот же небольшой угол θ.

Как будет выглядеть это же столкновение в других системах отсчета? Направим ось x вдоль биссектрисы угла θ и введем систему отсчета K, движущуюся вдоль оси x относительно системы центра масс со скоростью, равной x-составляющей скорости частицы *1*. В этой системе отсчета картина столкновения выглядит так, как показано на рис. 12.1, *б*: частица *1* движется параллельно оси y, изменив при столкновении направление скорости и импульса на противоположное. Сохранение y-составляющей полного импульса системы частиц при столкновении выражается соотношением

$$p_{1y}+p_{2y}=\tilde{p}_{1y}+\tilde{p}_{2y},$$

где $\tilde{p}_1$ и $\tilde{p}_2$ — импульсы частиц после столкновения. Так

как $\tilde{p}_{1y}=-p_{1y}$ и $\tilde{p}_{2y}=-p_{2y}$ (см. рис. 12.1, б), требование сохранения импульса означает равенство величин y-составляющих импульса частиц 1 и 2 в системе отсчета K:

$$p_{1y}=\tilde{p}_{2y}.$$

Теперь, наряду с K, введем в рассмотрение систему отсчета K', которая движется относительно системы центра масс со скоростью, равной x-составляющей скорости частицы 2. В этой системе K' частица 2 до и после столкновения движется параллельно оси y' (рис. 12.1, в). Применяя закон сохранения импульса, убеждаемся, что в этой системе отсчета, как и в системе K, имеет место равенство y-составляющих импульса частиц 1 и 2:

$$p'_{1y}=\tilde{p}'_{2y}.$$

Но из симметрии картин столкновения на рис. 12.1, б и в легко сделать вывод о том, что величина импульса частицы 1 в системе K равна величине импульса частицы 2 в системе отсчета K':

$$p_{1y}=\tilde{p}'_{2y}.$$

Сопоставляя два последних равенства, находим $p_{1y}=p'_{1y}$, т. е. y-составляющая импульса частицы 1 одинакова в системах отсчета K и K'. Точно так же находим $p_{2y}=p'_{2y}$. Другими словами, y-составляющая импульса любой частицы, перпендикулярная направлению относительной скорости систем отсчета, одинакова в этих системах.

Но y-составляющая скорости частицы имеет различное значение в системах отсчета K и K'. Согласно формулам преобразования скорости (11.7)

$$u'_{1y}=u_{1y}\sqrt{1-\frac{v^2}{c^2}},$$

где v есть скорость системы K' относительно K. Таким образом, в K' y-составляющая скорости частицы 1 меньше, чем в K. Это уменьшение y-составляющей скорости частицы 1 при переходе от K к K' непосредственно связано с релятивистским преобразованием времени: одинаковое в K и K' расстояние между штрихами A и B (рис. 12.1, б и в) частица 1 в системе K' проходит за большее время, чем в K. Если в K это время равно τ_0 (собственное время, так как оба события — пересечение штрихов A и B — происходят

в K при одном и том же значении координаты x), то в системе K' это время больше и равно $\tau=\tau_0/\sqrt{1-v^2/c^2}$.

Вспоминая теперь, что y-составляющая импульса частицы *1* одинакова в системах K и K', мы видим, что в системе K', где y-составляющая скорости частицы меньше, этой частице нужно приписать бо́льшую массу. Масса частицы, характеризующая пропорциональность между скоростью и импульсом, зависит от системы отсчета, т. е. является величиной относительной. В той системе отсчета, где скорость частицы много меньше скорости света, для связи между импульсом и скоростью можно написать обычное классическое выражение $\boldsymbol{p}=m_0\boldsymbol{u}$, где m_0 есть масса частицы в том смысле, как она понимается в нерелятивистской физике (масса покоя).

Будем считать, что в рассматриваемом нами «скользящем» столкновении скорость частицы *1* в системе K u_{1y} много меньше скорости света, т. е. ее масса в K есть масса покоя m_0, и $p_{1y}=m_0u_{1y}$. Написав аналогичное выражение для y-составляющей импульса в K', т. е. $p'_{1y}=mu'_{1y}$, видим, что равенство $p_{1y}=p'_{1y}$ будет обеспечено, если в K' приписать частице *1* бо́льшую массу m:

$$m=\frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}},$$

— уменьшение поперечной составляющей скорости при переходе от K к K' должно быть скомпенсировано возрастанием массы частицы. Точно так же в системе K' должны возрасти массы всех тел, неподвижных или движущихся с нерелятивистскими скоростями в лаборатории, связанной с системой K. Это возрастание массы тел, вызванное движением лаборатории, связано с релятивистским замедлением времени.

Масса частицы, движущейся со скоростью, близкой к скорости света, как видно из (12.1), может значительно превышать массу покоя. Если такой частице сообщить дополнительную энергию, чтобы увеличить ее импульс, то скорость ее при этом увеличится очень незначительно. Энергия частицы и ее импульс увеличиваются теперь за счет роста ее массы. Этот эффект наблюдается в работе ускорителей заряженных частиц высоких энергий. Экспериментальное подтверждение релятивистского возрастания

массы со скоростью служит, вероятно, лучшей проверкой теории относительности. Выясним теперь, к каким изменениям в выражении для энергии частицы приводит полученная выше формула для релятивистского импульса:

$$\boldsymbol{p} = m\boldsymbol{v}, \quad m = \frac{m_0}{\sqrt{1 - v^2/c^2}}. \qquad (12.2)$$

Для этого вспомним, что в механике приращение импульса частицы $\Delta \boldsymbol{p}$ за промежуток времени Δt равно импульсу действующей на нее силы $\boldsymbol{F}$:

$$\Delta \boldsymbol{p} = \boldsymbol{F} \Delta t.$$

Приращение кинетической энергии частицы $\Delta E_к$ за тот же промежуток времени равно работе силы $\boldsymbol{F}$:

$$\Delta E_к = \boldsymbol{F} \Delta \boldsymbol{r} = \boldsymbol{F}\boldsymbol{v} \Delta t = \boldsymbol{v} \Delta \boldsymbol{p} = \boldsymbol{v} \Delta (m\boldsymbol{v}). \qquad (12.3)$$

Здесь перемещение частицы $\Delta \boldsymbol{r}$ за Δt выражено через ее скорость $\boldsymbol{v}$. Из формулы (12.3) и будем исходить при получении выражения для релятивистской энергии. Перепишем формулу (12.2) для массы следующим образом:

$$m^2 (1 - v^2/c^2) = m_0^2.$$

Умножив обе части на c^2 и раскрыв скобки, получим

$$m^2c^2 - (m\boldsymbol{v})^2 = m_0^2c^2. \qquad (12.4)$$

При движении частицы под действием силы $\boldsymbol{F}$ ее скорость и, следовательно, релятивистская масса меняются. Рассмотрим приращения левой и правой частей (12.4) за промежуток времени Δt. Приращение правой части равно нулю, так как она не зависит от состояния движения частицы. Для нахождения приращения левой части воспользуемся тем, что приращение квадрата любой переменной величины f за малый промежуток времени приближенно равно

$$\Delta (f^2) = (f + \Delta f)^2 - f^2 \approx 2f \, \Delta f.$$

Поэтому

$$2mc \, \Delta (mc) - 2m\boldsymbol{v} \, \Delta (m\boldsymbol{v}) = 0.$$

При сокращении на $2m$ получаем

$$\Delta (mc^2) = \boldsymbol{v} \, \Delta (m\boldsymbol{v}). \qquad (12.5)$$

Правые части в выражениях (12.5) и (12.3) совпадают. Поэтому левая часть (12.5) есть приращение кинетической энергии частицы:

$$\Delta E_{\text{к}} = \Delta (mc^2). \qquad (12.6)$$

Мы получили, что приращение кинетической энергии частицы пропорционально приращению ее релятивистской массы. Так как кинетическая энергия покоящейся частицы равна нулю, а ее масса равна массе покоя, то из (12.6) находим

$$E_{\text{к}} = mc^2 - m_0c^2 = m_0c^2\left(\frac{1}{\sqrt{1-v^2/c^2}} - 1\right). \qquad (12.7)$$

Формула (12.7) дает выражение для релятивистской кинетической энергии частицы. Если скорость частицы мала по сравнению со скоростью света, формула (12.7) переходит в обычное выражение $E_{\text{к}} = m_0v^2/2$ для кинетической энергии частицы в нерелятивистской физике. Различие между классическим и релятивистским выражениями для кинетической энергии становится особенно существенным, когда скорость частицы приближается к скорости света. При $v \to c$ релятивистская кинетическая энергия (12.7) неограниченно возрастает: частица, обладающая конечной массой покоя m_0 и движущаяся со скоростью света, должна была бы иметь бесконечную кинетическую энергию. Зависимость кинетической энергии от скорости частицы показана на рис. 12.2.

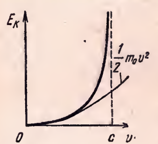


Рис. 12.2. Зависимость кинетической энергии тела от скорости.

Но вернемся к формуле (12.6), согласно которой приращение кинетической энергии тела сопровождается пропорциональным приращением его релятивистской массы. Вспомним, что важнейшим свойством энергии является ее способность превращаться из одной формы в другую в эквивалентных количествах при различных физических процессах — именно в этом заключается содержание закона сохранения и превращения энергии. Поэтому естественно ожидать, что возрастание релятивистской массы тела будет

происходить не только при сообщении ему кинетической энергии, но при любом увеличении энергии тела, независимо от конкретного вида энергии. Отсюда можно сделать фундаментальное заключение о том, что полная энергия

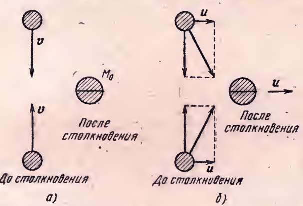


Рис. 12.3. Неупругое столкновение, наблюдаемое в разных системах отсчета.

тела пропорциональна его релятивистской массе независимо от того, из каких конкретных видов энергии она состоит.

Поясним сказанное на следующем простом примере. Рассмотрим неупругое столкновение двух одинаковых тел, движущихся навстречу друг другу с одинаковыми скоростями, так что в результате столкновения образуется одно тело, которое покоится (рис. 12.3, *а*). Пусть величина скорости каждого из тел до столкновения равна v, а масса покоя m_0. Массу покоя образовавшегося тела обозначим через M_0. Теперь рассмотрим это же столкновение с точки зрения наблюдателя в другой системе отсчета K', движущейся относительно исходной системы K налево (рис. 12.3, *б*) с малой (нерелятивистской) скоростью $-u$. Так как $u \ll c$, то для преобразования скорости при переходе от K к K' можно использовать классический закон сложения скоростей. Закон сохранения импульса требует, чтобы полный импульс тел до столкновения был равен импульсу образовавшегося тела. До столкновения полный импульс системы равен $\approx 2mu$, где $m = m_0/\sqrt{1-v^2/c^2}$ — релятивистская

вистская масса сталкивающихся тел, после столкновения $\approx M_0 u$, ибо вследствие $u \ll c$ массу образовавшегося тела и в K' можно считать равной массе покоя. Таким образом, из закона сохранения импульса следует, что масса покоя образовавшегося в результате неупругого соударения тела равна сумме релятивистских масс сталкивающихся частиц, т. е. она больше, чем сумма масс покоя исходных частиц:

$$M_0 = 2m = \frac{2m_0}{\sqrt{1-v^2/c^2}} > 2m_0. \tag{12.8}$$

Рассмотренный пример неупругого соударения двух тел, при котором происходит превращение кинетической энергии во внутреннюю энергию (в тепло), показывает, что увеличение внутренней энергии тела также сопровождается пропорциональным увеличением массы. Этот вывод должен быть распространен на все виды энергии: нагретое тело имеет бо́льшую массу, чем холодное, сжатая пружина имеет бо́льшую массу, чем несжатая, и т. п. Обобщение соотношения (12.6) на все виды энергии приводит нас к знаменитой формуле Эйнштейна:

$$E = mc^2 = \frac{m_0 c^2}{\sqrt{1-v^2/c^2}}. \tag{12.9}$$

Из (12.9) следует, что покоящееся тело обладает энергией $E_0 = m_0 c^2$. Эту энергию называют энергией покоя.

Закон пропорциональности массы и энергии является одним из самых значительных выводов теории относительности. Взаимосвязь массы и энергии заслуживает подробного обсуждения.

В механике масса тела есть физическая величина, являющаяся количественной характеристикой его инертных свойств, т. е. мера инертности. Это инертная масса. С другой стороны, масса характеризует способность тела создавать поле тяготения и испытывать силу в поле тяготения. Это тяготеющая, или гравитационная, масса. Инертность и способность к гравитационным взаимодействиям представляют собой совершенно различные проявления свойств материи. Однако то, что меры этих различных проявлений обозначаются одним и тем же словом, не случайно, а обусловлено тем, что оба свойства всегда существуют совместно и всегда друг другу пропорциональны, так что при надлежащем выборе единиц меры этих свойств можно выражать

одним и тем же числом. Равенство инертной и гравитационной масс есть экспериментальный факт, подтвержденный с огромной степенью точности в опытах Этвеша, Дикке и др. Как же следует отвечать на вопрос: есть ли масса инертная и масса гравитационная одно и то же или нет? По своим проявлениям они различны, но их численные характеристики пропорциональны друг другу. Такое положение вещей характеризуют словом «эквивалентность».

Аналогичный вопрос возникает в связи с понятиями массы и энергии в теории относительности. Проявления свойств материи, соответствующих массе и энергии, бесспорно различны. Но теория относительности утверждает, что эти свойства неразрывно связаны, а численные характеристики этих свойств пропорциональны друг другу. Поэтому в этом смысле можно говорить об эквивалентности массы и энергии. Всякое изменение энергии системы сопровождается эквивалентным изменением ее массы. Это относится как к изменениям кинетической энергии тела, при которых масса покоя остается неизменной, так и к изменениям различных видов внутренней энергии, при которых масса покоя меняется.

Опыт показывает нам, что в громадном большинстве физических процессов, в которых изменяется внутренняя энергия, масса покоя остается неизменной. Как это согласовать с законом пропорциональности массы и энергии? Дело в том, что обычно подавляющая часть внутренней энергии (и соответствующей ей массы покоя) в превращениях не участвует и в результате оказывается, что определяемая из взвешивания масса практически сохраняется, несмотря на то, что тело выделяет или поглощает энергию. Это объясняется просто недостаточной точностью взвешивания. Для иллюстрации рассмотрим несколько численных примеров.

1. Энергия, высвобождающаяся при сгорании нефти, при взрыве динамита и при других химических превращениях, представляется нам в масштабах повседневного опыта громадной. Однако если перевести ее величину на язык эквивалентной массы, то окажется, что эта масса не составляет и 10^{-10} от полной величины массы покоя. Например, при соединении 1 г водорода с 8 г кислорода выделяется около 10^{12} эрг энергии. Масса покоя образовавшейся воды на $\Delta m = \Delta E/c^2 \approx 10^{-9}$ г меньше массы исходных веществ.

Такое изменение массы слишком мало для того, чтобы его можно было обнаружить с помощью современных приборов.

2. При неупругом столкновении двух частиц по 1 г, разогнанных навстречу друг другу до скорости 1 км/с, добавочная масса покоя слипшейся пары составляет

$$\Delta m \approx \left(2\cdot\frac{mv^2}{2}\right): c^2 \approx 10^{-11} \text{ г}.$$

(При такой скорости можно пользоваться нерелятивистским выражением для кинетической энергии.) Эта величина намного меньше ошибки, с которой может быть измерена масса 1 г.

Естественно задать вопрос: почему при обычных условиях подавляющая часть энергии находится в совершенно пассивном состоянии и в превращениях не участвует? На этот вопрос теория относительности не может дать ответа. Ответ следует искать в области квантовых закономерностей, одной из характерных особенностей которых является существование устойчивых состояний с дискретными уровнями энергии. Для элементарных частиц энергия, соответствующая массе покоя, либо превращается в активную форму (излучение) целиком, либо вовсе не превращается. Примером может служить превращение пары электрон — позитрон в гамма-излучение. У атомов подавляющая часть массы находится в форме массы покоя элементарных частиц, которая в химических реакциях не изменяется. Даже в ядерных реакциях энергия, соответствующая массе покоя тяжелых частиц (нуклонов), входящих в состав ядер, остается пассивной. Но здесь активная часть энергии, т. е. энергия взаимодействия нуклонов, составляет уже заметную долю энергии покоя.

Таким образом, экспериментальное подтверждение релятивистского закона пропорциональности энергии и массы следует искать в мире физики элементарных частиц и ядерной физики. Рассмотрим в качестве примера одну из первых ядерных реакций, вызванных полученными на ускорителе протонами: превращение ядра лития в две альфа-частицы:

$${}_3\mathrm{Li}^7 + {}_1\mathrm{H}^1 \rightarrow 2\,{}_2\mathrm{He}^4.$$

Значение массы покоя атомных ядер может быть с высокой точностью определено с помощью масс-спектрометра. Так, масса покоя протона ${}_1\mathrm{H}^1$ равна 1,00783 атомной единицы массы (а. е. м.), ядра ${}_3\mathrm{Li}^7$ 7,01601 а. е. м., а α-частицы

($_2He^4$) 4,00260 а. е. м. Сравним полные массы покоя исходных ядер и продуктов реакции. Масса покоя ядер, вступающих в реакцию, равиа 8,02384 а. е. м., а масса покоя коиечиых продуктов меиьше: 8,00520 а. е. м. Таким образом, в результате реакции масса покоя уменьшается на величииу $\Delta m = 0{,}01864$ а. е. м. Соответствующая этому изменеиию массы энергия $\Delta mc^2 = 17{,}4$ МэВ с хорошей точностью совпадает с измереиной на опыте киеитической эиергией образующихся α-частиц. (Первоиачальиая киие-тическая эиергия протона мала по сравиеиию с этой величииой, и поэтому при расчете ее можио ие прииимать во виимаиие.)

§ 13. Примеры релятивистского движения частиц

Прежде всего получим формулу, связывающую энергию и импульс релятивистских частиц. Напомиим релятивистские формулы для эиергии и импульса частиц (12.9) и (12.2):

$$E = \frac{m_0 c^2}{\sqrt{1 - v^2/c^2}}, \tag{13.1}$$

$$\boldsymbol{p} = \frac{m_0 \boldsymbol{v}}{\sqrt{1 - v^2/c^2}}. \tag{13.2}$$

Сравиивая (13.1) и (13.2), получаем простую формулу, связывающую энергию, импульс и скорость частицы:

$$\boldsymbol{p} = \frac{E}{c^2} \boldsymbol{v}. \tag{13.3}$$

Возводим обе части (13.1) в квадрат и записываем в виде

$$E^2 - E^2 \frac{v^2}{c^2} = m_0^2 c^4.$$

Замеияя, согласио (13.3), во втором слагаемом в левой части $E^2 v^2$ на $p^2 c^4$, иаходим

$$E^2 - p^2 c^2 = m_0^2 c^4. \tag{13.4}$$

Полученное соотиошеиие (13.4), связывающее эиергию и импульс частицы, является одной из важиейших формул релятивистской физики. Обратим внимание иа то, что в правой части этой формулы стоит величина, не зависящая от выбора системы отсчета. Поэтому, хотя каждое из слагаемых в левой части имеет разное значение в различных

инерциальных системах отсчета, вся левая часть не зависит от выбора системы отсчета, т. е. представляет собой релятивистский инвариант.

В качестве первого примера рассмотрим движение первоначально покоившейся частицы с зарядом q и массой покоя m_0 в однородном электрическом поле напряженности $\boldsymbol{E}$. Действующая на частицу сила $\boldsymbol{F}$ постоянна и равна $q\boldsymbol{E}$. Поэтому из закона изменения импульса $\Delta\boldsymbol{p}=\boldsymbol{F}\Delta t$ немедленно следует, что

$$\boldsymbol{p}=\boldsymbol{F}t. \tag{13.5}$$

Подставляя это выражение для импульса частицы в формулу (13.4), получим

$$E^2=(Ft)^2c^2+m_0^2c^4. \tag{13.6}$$

Теперь с помощью (13.3) находим скорость частицы $\boldsymbol{v}$ спустя промежуток времени t после начала движения:

$$\boldsymbol{v}=\frac{Ft}{m_0\sqrt{1+(Ft/m_0c)^2}}. \tag{13.7}$$

Если $Ft/m_0c \ll 1$, т. е. электрическое поле слабое или мало́ время движения, то в подкоренном выражении в (13.7) можно пренебречь вторым слагаемым, и для скорости получается обычное нерелятивистское выражение:

$$\boldsymbol{v}=\frac{F}{m_0}t.$$

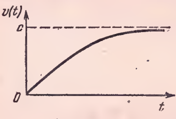


Рис. 13.1. Скорость частицы при движении в однородном электрическом поле.

Если $Ft/m_0c \gg 1$, то под корнем в (13.7) можно пренебречь единицей по сравнению со вторым слагаемым. Видно, что при этом величина скорости v стремится к c. На рис. 13.1 показана зависимость скорости v от времени.

Перейдем к рассмотрению движения заряженной частицы в однородном магнитном поле. Поскольку действующая на частицу со стороны магнитного поля сила Лоренца перпендикулярна скорости частицы, то скорость не меняется по величине и, следовательно, не меняется и масса частицы m. Поэтому закон изменения импульса частицы

запишется в виде

$$m\frac{d\boldsymbol{v}}{dt}=q\boldsymbol{v}\times\boldsymbol{B}. \qquad (13.8)$$

Если скорость $\boldsymbol{v}$ перпендикулярна вектору индукции магнитного поля $\boldsymbol{B}$, то частица движется по окружности и величина ее ускорения равна v^2/R, где R — радиус окружности. В этом случае уравнение (13.8) дает

$$m\frac{v^2}{R}=qvB. \qquad (13.9)$$

Для угловой скорости вращения ω_c, связанной с линейной скоростью v обычным соотношением $v=\omega_c R$, с помощью (13.9) находим

$$\omega_c=\frac{qB}{m}. \qquad (13.10)$$

Выражение (13.10) имеет точно такой же вид, как и нерелятивистская формула для угловой скорости вращения в магнитном поле, только в знаменателе стоит релятивистская масса частицы m, связанная с ее массой покоя m_0 соотношением $m=m_0/\sqrt{1-v^2/c^2}$.

В качестве третьего примера рассмотрим ускоритель заряженных частиц на встречных пучках. Выясним, в чем преимущество таких ускорителей по сравнению с обычными ускорителями с неподвижной мишенью, и установим соответствие между кинетической энергией частицы $E_к$ в обычном ускорителе и эквивалентной энергией $E'_к$ в ускорителе со встречными пучками.

Одной из важнейших характеристик ускорителя является та доля кинетической энергии разогнанных элементарных частиц, которая может быть использована для реакции образования новых частиц. В обычных ускорителях, когда частица-мишень неподвижна, требование сохранения импульса исключает возможность превращения всей кинетической энергии частицы-снаряда в энергетический эквивалент массы покоя новых частиц, образующихся при столкновении. Поскольку до столкновения суммарный импульс снаряда и мишени отличен от нуля, то такой же суммарный импульс должен быть и после столкновения. Поэтому образовавшиеся в результате столкновения частицы не могут находиться в покое и, следовательно, часть

начальной кинетической энергии снаряда переходит в кинетическую энергию частиц после столкновения.

Однако если сталкивающиеся частицы с равными массами летят навстречу друг другу с одинаковыми скоростями, то в результате неупругого удара вся кинетическая энергия налетающих частиц может быть использована для рождения новых частиц: поскольку начальный импульс

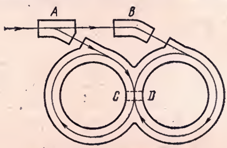


Рис. 13.2. Накопительные кольца.

системы равен нулю, то ничто не запрещает покоиться образовавшимся в результате столкновения частицам.

Оценим вначале «выигрыш» в энергии для простого случая столкновения одинаковых нерелятивистских частиц. Используя закон сохранения импульса, легко убедиться, что в этом случае при неподвижной мишени для реакции образования новых частиц может быть использована только половина кинетической энергии налетающей частицы $E_к$: $\Delta E = E_к/2$. Если же столкнутся движущиеся навстречу друг другу частицы с кинетическими энергиями $E_к$, то для реакции может быть использована вся их кинетическая энергия: $\Delta E = 2E_к$. Таким образом, используя ускоритель, способный разогнать частицы до кинетической энергии $E_к$, мы можем с помощью накопительных колец повысить эффективность использования кинетической энергии в 4 раза.

Идея устройства накопительных колец показана на рис. 13.2. Пучок частиц из ускорителя с помощью быстродействующего магнита-переключателя A разделяется на два пучка, каждый из которых с помощью системы отклоняющих магнитов A и B направляется в свое кольцо, где обращается по орбите благодаря удерживающему магнитному полю, перпендикулярному плоскости рисунка. На

общем участке CD происходят столкновения движущихся навстречу друг другу частиц.

Итак, в нерелятивистском случае неупругого столкновения частиц одинаковой массы, одна из которых покоится, т. е. при использовании неподвижной мишени, только половина первоначальной энергии может перейти в энергию покоя рождающихся частиц. А как обстоит дело в случае релятивистских частиц, с которыми имеет дело физика высоких энергий? Оказывается, что для неподвижной мишени еще хуже. Чтобы убедиться в этом, придется тщательно рассмотреть законы сохранения энергии и импульса при столкновении релятивистских частиц.

Рассмотрим неупругое столкновение релятивистской частицы с массой покоя m_0 с такой же покоящейся частицей.

Будем искать энергию ΔE, которая может быть использована для образования новых частиц в этом случае. Обозначим через M_0 полную массу покоя системы после столкновения. Тогда ΔE есть не что иное, как увеличение энергии покоя частиц, которое произошло в рассматриваемом столкновении:

$$\Delta E = M_0c^2 - 2m_0c^2. \tag{13.11}$$

Найдем теперь M_0 — массу покоя частиц системы после столкновения. Применим к столкновению законы сохранения энергии и импульса. Из формулы (13.4) выразим квадрат импульса любой частицы p^2 через ее полную энергию E:

$$p^2 = \frac{E^2}{c^2} - m_0^2c^2. \tag{13.12}$$

Полная энергия релятивистской частицы E есть сумма энергии покоя частицы и ее кинетической энергии:

$$E = m_0c^2 + E_к. \tag{13.13}$$

Энергия, которой характеризуют ускорители,— это кинетическая энергия разогнанных частиц $E_к$. Учитывая, что до столкновения одна из частиц покоилась ($p_2=0$), запишем квадрат импульса всей системы до удара P^2, равный квадрату импульса налетающей частицы p_1^2 (13.12), в виде

$$P^2 = \left(\frac{m_0c^2 + E_к}{c}\right)^2 - m_0^2c^2. \tag{13.14}$$

Согласно закону сохранения энергии полная энергия системы после столкновения E' такая же, как и до столкновения, т. е. равна сумме энергий покоя обеих частиц и кинетической энергии налетающей частицы:

$$E' = 2m_0c^2 + E_к. \tag{13.15}$$

Запишем теперь квадрат импульса системы после столкновения с помощью (13.12) и (13.15):

$$P^2 = \frac{E'^2}{c^2} - M_0c^2 = \left(\frac{2m_0c^2 + E_к}{c}\right)^2 - M_0c^2. \tag{13.16}$$

Полный импульс системы до удара (13.14) и после удара (13.16) обозначены одной и той же буквой P, так как полный импульс системы сохраняется. Приравнивая правые части равенств (13.14) и (13.16), после простых преобразований находим M_0:

$$M_0 = 2m_0\left(1 + \frac{E_к}{2m_0c^2}\right)^{1/2}. \tag{13.17}$$

Теперь для ΔE в (13.11) получим

$$\Delta E = 2m_0c^2\left(\sqrt{1 + \frac{E_к}{2m_0c^2}} - 1\right). \tag{13.18}$$

Легко видеть, что для нерелятивистской частицы, кинетическая энергия которой много меньше энергии покоя: $E_к \ll m_0c^2$, выражение (13.18) дает результат, полученный ранее элементарным путем: $\Delta E = E_к/2$. Для этого достаточно воспользоваться приближенной формулой

$$\sqrt{1+x} \approx 1 + x/2 \quad \text{при } x \ll 1.$$

В противоположном ультрарелятивистском случае, когда кинетическая энергия частицы много больше энергии покоя: $E_к \gg m_0c^2$, в формуле (13.18) можно пренебречь единицами по сравнению с $E_к/(2m_0c^2)$. Тогда выражение (13.18) принимает вид

$$\Delta E \approx \sqrt{E_к \cdot 2m_0c^2}. \tag{13.19}$$

Если, например, мы хотим иметь $\Delta E = 20$ ГэВ при столкновении протонов (энергия покоя протона $m_0c^2 \approx 1$ ГэВ), то с помощью формулы (13.19) убеждаемся, что необходим ускоритель, разгоняющий протоны до энергии $E_к \approx 200$ ГэВ. Таким образом, в рассматриваемом примере может быть

использована только десятая часть кинетической энергии протона (а не половина, как было бы в нерелятивистском случае). Итак, из-за релятивистских эффектов доля кинетической энергии разогнанных частиц, которая может быть использована для реакции, у ускорителей с неподвижной мишенью падает с ростом энергии. В ускорителе же на встречных пучках и в релятивистском случае вся кинетическая энергия сталкивающихся частиц может перейти в энергию покоя рождающихся частиц.

Интересно получить соотношение, связывающее кинетические энергии частиц в ускорителе обычного типа $E_к$ и ускорителе на встречных пучках $E'_к$, при которых получается одна и та же энергия ΔE, способная превратиться в энергию покоя рождающихся частиц. В ускорителе на встречных пучках $\Delta E = 2E'_к$. В ускорителе с неподвижной мишенью ΔE определяется формулой (13.18). Подставляя в нее $\Delta E = 2E'_к$, находим

$$E_к = 2m_0c^2\left[\left(\frac{E'_к}{m_0c^2}+1\right)^2-1\right]. \qquad (13.20)$$

В ультрарелятивистском случае, когда $E'_к \gg m_0c^2$, эта формула принимает вид

$$E_к \approx \frac{2E_к'^2}{m_0c^2}. \qquad (13.21)$$

Из приведенных формул видно, что выигрыш при использовании ускорителей на встречных пучках особенно велик для легких частиц, например электронов, для которых $m_0c^2 \approx 0{,}5$ МэВ. Так, для установки со встречными пучками, ускоряющей электроны до энергии $E'_к = 130$ МэВ, энергия $E_к$ эквивалентного ускорителя с неподвижной мишенью составляет, согласно формуле (13.21), примерно 70 ГэВ, т. е. в 520 раз больше!

§ 14. Принцип эквивалентности. Гравитационное смещение спектральных линий

Заканчивая знакомство с основами теории относительности, остановимся коротко на основных физических идеях, которые легли в основу созданной Эйнштейном релятивистской теории тяготения. Математическая сторона этой теории очень сложна, однако ее физические основы весьма просты и могут быть изложены в доступной форме.

Релятивистская теория тяготения основана на экспериментальном факте равенства инертной и гравитационной масс, который установлен с исключительно высокой степенью точности. В классической физике это равенство носило случайный характер. Основываясь на равенстве инертной и гравитационной масс, Эйнштейн сформулировал так называемый принцип эквивалентности, послуживший отправным пунктом при создании релятивистской теории тяготения.

Для того чтобы легче было понять содержание принципа эквивалентности, рассмотрим следующий мысленный опыт. Пусть в однородном поле тяжести напряженности $\boldsymbol{g}$ неподвижно висит лифт. Все тела в лифте подвергаются действию земного поля тяготения и, предоставленные самим себе, свободно падают относительно лифта с одним и тем же ускорением $\boldsymbol{g}$. Вообразим теперь, что лифт настолько удален от Земли и прочих небесных тел, что он практически не подвергается с их стороны никаким гравитационным воздействиям. Но пусть зато кто-то тянет за трос лифта, сообщая ему постоянное ускорение $\boldsymbol{a}=-\boldsymbol{g}$, направленное «вверх». Что будет происходить со свободными телами в таком поступательно движущемся лифте? Так как гравитационного поля в лифте нет, то в инерциальной системе отсчета первоначально покоившиеся тела будут продолжать покоиться. Однако относительно ускоренно движущегося лифта эти тела будут двигаться с одинаковым ускорением $\boldsymbol{g}$. Таким образом, наблюдатель, находящийся в закрытом лифте и не имеющий возможности «выглянуть наружу», по поведению свободных тел не сможет решить, покоится ли лифт в однородном поле тяжести напряженности $\boldsymbol{g}$ или движется с ускорением $\boldsymbol{a}=-\boldsymbol{g}$ в отсутствие гравитационного поля под действием каких-то внешних сил. Оказывается, что такая эквивалентность поля тяжести и ускоренного движения системы отсчета имеет место для любых механических явлений: все механические явления в движущемся с ускорением лифте будут в точности такими же, как и в неподвижном лифте, но висящем в поле тяжести.

Эйнштейн предположил, что это утверждение справедливо не только для механических, но и для любых физических явлений вообще: все физические явления в равномерно ускоренном лифте будут происходить точно так же, как и в неподвижном лифте, висящем в однородном поле

тяжести. Это и есть принцип эквивалентности. Ему можно придать несколько иной вид. Предположим, что внешние силы, сообщающие лифту ускорение a в отсутствие гравитационного поля, исчезли. Тогда система отсчета, связанная с лифтом, будет инерциальной. Но эквивалентность поля тяжести и ускоренного движения системы отсчета под действием внешних сил означает, что и в свободно падающем в поле тяжести лифте законы физики такие же, как и в инерциальной системе отсчета. Другими словами, ускорение системы отсчета, свободно падающей в поле тяжести, полностью компенсирует действие силы тяжести. Это связано с тем, что всем без исключения телам гравитационное поле сообщает одно и то же ускорение.

Дальнейшее математическое развитие принципа эквивалентности приводит к релятивистской теории тяготения. Эта теория пришла на смену ньютоновской теории тяготения, но в той области, где ньютоновская теория блестяще выдержала многочисленные проверки на опыте, эйнштейновская теория приводит к тем же результатам. Для всех практических нужд, например для космических полетов и расчетов движения планет и других небесных тел, ньютоновская теория является вполне достаточной. Однако в специально проведенных экспериментах удалось наблюдать исключительно тонкие явления, которые можно объяснить только с помощью релятивистской теории тяготения.

Принцип эквивалентности может с успехом использоваться и при анализе обычных физических явлений. Часто его применение позволяет очень просто решить задачи, исследование которых без использования принципа эквивалентности представляет значительные трудности. Проиллюстрируем это на следующем примере. Пусть на горизонтальной плоскости находится закрытый сосуд кубической формы, наполовину заполненный водой. В некоторый момент времени сосуд начинает двигаться в горизонтальном направлении с ускорением $a=g$ (рис. 14.1, a). Выясним, на сколько градусов при этом нагреется вода в сосуде.

То, что вода в сосуде действительно должна нагреваться, по-видимому, не вызывает сомнения: ведь при начальном толчке возникают колебания воды, которые постепенно благодаря трению о стенки и вязкости воды затухают. Но как подсчитать выделившуюся энергию?

Попытка «в лоб» применить закон сохранения энергии обречена на неудачу: чем больше мы будем об этом думать, тем очевиднее станет безнадежность этой затеи. Действительно, мы не только не сможем подсчитать работу внешней силы, разгоняющей сосуд, мы не сможем даже найти

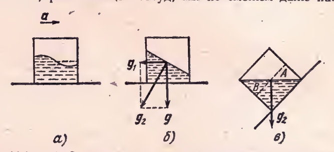


Рис. 14.1. *а*) Сосуд с водой начинает движение с ускорением; *б*) положение воды в сосуде после прекращения колебаний; *в*) к вычислению изменения потенциальной энергии воды в эффективном поле тяжести.

величину этой силы до тех пор, пока не прекратятся колебания воды в сосуде.

Попробуем применить в этой задаче принцип эквивалентности. Вместо того, чтобы рассматривать ускоренно движущийся сосуд, будем считать, что сосуд неподвижен, но на воду в нем действует дополнительное «гравитационное» поле напряженности $\boldsymbol{g}_1 = -\boldsymbol{a}$ (рис. 14.1, *б*). Это поле, складываясь с истинным полем тяжести Земли, дает эффективное поле тяжести, напряженность которого $\boldsymbol{g}_2 = \boldsymbol{g} + \boldsymbol{g}_1 = \boldsymbol{g} - \boldsymbol{a}$. Теперь осталось только представить себе начальное и конечное положение воды в сосуде в этом эффективном поле тяжести. В начальный момент поверхность воды в сосуде была горизонтальной, т. е. по отношению к вектору напряженности эффективного поля тяжести $\boldsymbol{g}_2$ она занимала наклонное неравновесное положение, указанное пунктиром на рис. 14.1, *в*. Затем возникли колебания жидкости, в процессе которых происходили многократные превращения потенциальной энергии в кинетическую и обратно. Благодаря трению кинетическая энергия воды постепенно переходила во внутреннюю энергию, и в конце концов вода успокоилась в новом равновесном положении, в котором ее поверхность перпендикулярна вектору $\boldsymbol{g}_2$.

Конечное положение уровня воды в сосуде для рассматриваемого случая $a=g$ показано на рис. 14.1, *в*. На основании закона сохранения энергии можно утверждать, что увеличение внутренней энергии воды равно убыли ее потенциальной энергии в эффективном поле тяжести g_2 при переходе из начального состояния в конечное. Из рис. 14.1, *в* видно, что перемещение воды в сосуде в конечном счете свелось к тому, что часть жидкости в объеме, обозначенном A, перешла в положение B. Теперь убыль потенциальной энергии воды вычисляется элементарно. Учитывая, что $g_2=\sqrt{2}\,g$, найдем

$$E_1-E_2=\frac{g\rho l^4}{24},$$

где ρ — плотность воды, l — длина ребра куба. Изменение температуры воды ΔT найдем, разделив увеличение ее внутренней энергии, равное E_1-E_2, на теплоемкость всей массы жидкости:

$$\Delta T=\frac{gl}{12c_в},$$

где $c_в$— удельная теплоемкость воды. Интересно отметить, что изменение температуры воды зависит от размеров куба, но для разумных размеров сосуда ($l\approx 1$ м) это изменение ничтожно.

Применимость принципа эквивалентности не только к механическим, но и к любым физическим явлениям вообще можно проиллюстрировать на примере гравитационного смещения спектральных линий, которое было в начале шестидесятых годов зарегистрировано в лабораторных условиях на Земле. Источник монохроматического гамма-излучения располагался у поверхности Земли, а приемник — на высоте $H=22$ м над источником. Частота регистрируемого приемником излучения была сдвинута в «красную» сторону, т. е. в сторону меньших частот по сравнению с частотой источника. Чем может быть вызвано изменение частоты электромагнитной монохроматической волны? Вспомним эффект Доплера: при относительном движении источника монохроматического излучения с частотой ν_0 и приемника происходит сдвиг частоты. Он определяется соотношением (см. формулу (1.16), стр. 529)

$$\frac{\Delta\nu}{\nu_0}=\frac{v}{c}. \qquad (14.1)$$

Эта формула написана для случая, когда $v \ll c$ и движение происходит по прямой, соединяющей источник с приемником. В (14.1) скорость v нужно считать положительной при сближении источника и приемника и отрицательной при удалении. Но как же использовать это явление в рассматриваемом случае — ведь источник и приемник неподвижны? Воспользуемся принципом эквивалентности: наличие однородного гравитационного поля напряженности $\boldsymbol{g}$ в инерциальной системе отсчета эквивалентно ускоренному движению системы отсчета с ускорением $-\boldsymbol{g}$ в отсутствие гравитационного поля. Применительно к нашей задаче это означает, что можно «забыть» о поле тяготения, но считать, что источник и приемник движутся с ускорением $\boldsymbol{a}=-\boldsymbol{g}$, которое направлено вверх. Если считать, что излучение волны с частотой ν_0 происходит в тот момент, когда скорость источника равна нулю, то спустя время $\Delta t=H/c$, когда волна достигнет приемника, его скорость будет равна $g\,\Delta t=gH/c$. При вычислении относительной скорости v, входящей в формулу (14.1), скорость источника следует брать в момент излучения, а скорость приемника — в момент прихода волны. Поэтому использование формулы (14.1) немедленно показывает, что вследствие эффекта Допплера будет наблюдаться сдвиг частоты, равный

$$\frac{\Delta\nu}{\nu_0}=-\frac{gH}{c^2}. \qquad (14.2)$$

Используя значения $g\approx 10$ м/с², $H=22$ м, $c=3\cdot 10^8$ м/с, видим, что $\Delta\nu/\nu_0\approx -2{,}4\cdot 10^{-15}$. Знак минус означает, что частота γ-излучения уменьшается, т. е. в данных условиях, когда приемник находится выше источника, наблюдается «красное» гравитационное смещение. Если поменять местами источник и приемник, то частота увеличится. Замечательно, что, несмотря на столь малую величину эффекта, его удалось не только обнаружить на опыте, но и измерить с точностью до нескольких процентов! Метод, позволивший наблюдать столь ничтожный сдвиг частоты, основан на использовании эффекта Мёссбауэра. Чтобы лучше представить себе чувствительность этого метода, позволившего зарегистрировать относительное изменение частоты, равное 10^{-15}, отметим, что это эквивалентно возможности заметить изменение массы тела в миллион тонн при добавлении к нему одного миллиграмма!

Явление гравитационного смещения спектральных линий можно, разумеется, объяснить и не прибегая к принципу эквивалентности, а используя законы сохранения энергии или импульса. Хотя постоянная Планка h не входит в выражение (14.2) для относительного сдвига частоты и, следовательно, эффект гравитационного смещения не связан с квантовой природой излучения, для применения законов сохранения удобно представить гамма-излучение как поток фотонов или гамма-квантов с энергией $h\nu$ и импульсом $h\nu/c$.

Применим к процессу распространения гамма-квантов в гравитационном поле закон сохранения энергии. Каков бы ни был характер взаимодействия γ-квантов с гравитационным полем, полная энергия системы остается неизменной, ибо поле тяготения потенциально. Пусть γ-квант, испускаемый источником на поверхности Земли, имеет частоту ν_0. Тогда энергия каждого γ-кванта есть $E_0=h\nu_0$. Гамма-квант, зарегистрированный приемником на высоте H, будет иметь другую частоту: $\nu_1\neq\nu_0$, ибо при подъеме в поле тяжести потенциальная энергия этого кванта увеличилась на величину mgH, где m — масса γ-кванта, определяемая соотношением $m=h\nu/c^2$. На основании закона сохранения энергии

$$h\nu_0=h\nu_1+mgH. \qquad (14.3)$$

Легко убедиться, что изменение потенциальной энергии γ-кванта много меньше $h\nu_0$, ибо $mgH/(h\nu)=gH/c^2\ll 1$. Поэтому в выражении (14.3) массу γ-кванта можно считать постоянной и равной $h\nu_0/c^2$. Строго говоря, если бы масса не считалась постоянной, изменение потенциальной энергии нельзя было бы полагать равным mgH. Подставляя значение массы в (14.3), получим

$$h\nu_0=h\nu_1+\frac{h\nu_0}{c^2}gH, \qquad (14.4)$$

откуда для относительного изменения частоты получаем выражение (14.2).

Этот же результат можно получить и с помощью закона сохранения импульса. Для этого надо проследить, как и почему меняется импульс γ-кванта при его движении от источника до приемника. Начальный импульс $p_0=h\nu_0/c$, конечный $p_1=h\nu_1/c$. Поскольку на γ-квант во время

движения от источника до приемника $\Delta t = H/c$ действует сила тяжести $mg = (h\nu_0/c^2)\,g$, то закон изменения импульса $p_1 - p_0 = -mg\,\Delta t$ запишется следующим образом:

$$\frac{h}{c}\,\Delta\nu = -\frac{h\nu_0}{c^2}\,g\,\frac{H}{c},$$

откуда немедленно следует соотношение (14.2).

Таким образом, смещение спектральных линий возникает из-за изменения энергии и импульса γ-кванта под действием тяготения. Разумеется, эффект гравитационного смещения будет наблюдаться и в неоднородном поле тяготения. Проще всего это рассмотреть с помощью подхода, основанного на законе сохранения энергии. Нетрудно сообразить, как нужно видоизменить формулу (14.3), чтобы рассмотреть случай неоднородного поля. С таким случаем мы сталкиваемся, например, при наблюдении на Земле спектра излучения, испускаемого атомами на звезде или на Солнце. Так, гравитационное смещение частоты, вызываемое полем Солнца, составляет $\Delta\nu/\nu_0 \approx -2\cdot 10^{-6}$, что в миллиард раз больше, чем зафиксированное на опыте гравитационное смещение частоты γ-квантов в лабораторных условиях на Земле. Однако наблюдение гравитационного красного смещения спектральных линий Солнца связано с большими трудностями, возникающими из-за наличия допплеровского смещения этих линий, обусловленного движением газа в фотосфере Солнца.

6. ВВЕДЕНИЕ В КВАНТОВУЮ ФИЗИКУ *)

ЗАКОНЫ МИКРОМИРА

§ 1. Световые кванты

К началу XX века в физике накопился ряд экспериментальных фактов, не допускавших объяснения в рамках той теории, которая называется теперь классической физикой. Обобщение данных различных опытов привело к следующим не известным ранее важным выводам: во-первых, к представлению о двойственном характере электромагнитного излучения, проявляющего то волновые, то корпускулярные свойства, и, во-вторых, к утверждению о существовании дискретных значений некоторых из тех физических величин, которые, по представлениям классической физики, могли меняться непрерывно.

Начнем с обсуждения явления фотоэлектрического эффекта. Опыты показали, что кинетическая энергия электронов, испускаемых поверхностью металла, освещаемой видимым или ультрафиолетовым светом, не зависит от интенсивности излучения, а зависит лишь от рода металла и от частоты излучения ν по следующему закону:

$$\frac{1}{2}mv^2 = h\nu - A. \qquad (1.1)$$

В этом выражении A — так называемая работа выхода электрона из металла, т. е. энергия, необходимая для удаления электрона из этого металла. В 1905 году Эйнштейн связал этот результат с гипотезой Планка о том, что свет излучается отдельными порциями, причем энергия каждой такой порции определяется формулой $E = h\nu$.

---

*) В этом разделе, посвященном атомной физике, используется система единиц СГСЭ, как это повсеместно принято.

Соотношение (1.1) получило название уравнения Эйнштейна. Детальное изучение опытных данных показало, что h в выражении (1.1) совпадает с универсальной постоянной, входящей в теорию Планка: $h=6{,}626\cdot 10^{-27}$ эрг·с.

Попытаемся понять, почему электрон поглощает только один квант энергии независимо от интенсивности излучения. Самое простое объяснение этого явления заключается в следующем: свет состоит из частиц, т. е. из излученных порций световой энергии $E=h\nu$, которые сохраняют свою индивидуальность в процессе распространения и в дальнейшем, при столкновении с электроном, передают ему всю свою энергию.

Это предположение подкрепляется различными опытами, например такими, когда на фотопластинку направляются лучи очень низкой интенсивности. На пластинке получаются хаотически расположенные темные пятна со средней плотностью, пропорциональной интенсивности света. Для очень интенсивного пучка распределение пятен становится настолько плотным, что они практически непрерывны. При такой интенсивности пучка света он становится эквивалентным тому, что в классической физике называется световой волной.

Хотя предположение, что свет состоит из локализованных частиц, позволяет просто объяснить фотоэлектрический эффект, оно не согласуется с огромной совокупностью экспериментов, приводящих к выводу, что свет является формой волнового движения. Вспомним чередование интенсивности дифракционных полос света, падающего на экран после прохождения через одну или несколько щелей. Часто имеет место такое явление, что при двух открытых щелях, близких друг к другу, интенсивность света в некоторых местах экрана будет очень мала, в то время как в тех же местах при пропускании света только через одну щель наблюдается высокая интенсивность. Эти результаты легко объясняются предположением, что свет представляет собой волны, которые, интерферируя, могут усиливаться или ослабляться. Но их совершенно невозможно объяснить, если предположить, что свет состоит из локализованных частиц. Такие частицы должны были бы проходить или через одну, или через другую щель, и наличие второй щели едва ли могло бы влиять на характер движения частиц, проходящих через первую щель.

Но, может быть, возможно объяснить закономерности фотоэффекта, исходя из волновых представлений о свете? Попробуем рассмотреть фотоэффект с классической точки зрения. Взаимодействуя с электроном, находящимся внутри атома, излучение передает ему свою энергию. Электрон будет поглощать энергию световой волны, пока он не освободится из атома. Объяснить уравнение (1.1) можно, предположив, что свойства атома таковы, что электрон будет сохранять полученную от света энергию и находиться в атоме до тех пор, пока не накопит ее до величины $h\nu$, после чего он покидает атом. Если бы атом действительно обладал такими свойствами, то для света с очень маленькой интенсивностью фотоэффект не должен был бы наблюдаться в течение очень долгого времени, так как должно было бы пройти значительное время, чтобы накопилась необходимая порция энергии. Соответствующие опыты проводились с металлическими пылинками и очень слабым светом. Пылинки были настолько малы, что потребовалось бы много часов для накопления энергии $h\nu$. Однако немедленно после начала их освещения появлялось некоторое количество фотоэлектронов. Итак, эта попытка объяснения не удалась.

Такова же была судьба всех остальных подобных попыток объяснения закономерностей фотоэффекта на основе волновых представлений. Это означает, что волновая теория не способна объяснить внезапную локализацию конечных порций энергии на одной частице.

Итак, одна группа опытов указывает, что свет — это частицы, которые могут быть локализованы; другая же группа не менее убедительных опытов доказывает, что свет — это волны. Какое же из этих утверждений правильно?

Положение напоминает притчу о слепом и глухом путниках, застигнутых грозой: для глухого молния — только яркая вспышка света, а для слепого — только раскаты грома. Вопрос стоит так: можно ли найти единое представление, объясняющее все наши сведения о природе света, так же как наше понятие о молнии объединяет представления о ней слепого и глухого путников? Позже мы вернемся к обсуждению этого вопроса, а пока будем просто считать, что свет существует в форме квантов-фотонов, которые при некоторых условиях проявляют себя подобно частицам, а при других — подобно волнам.

Мы уже подробно рассматривали волновые свойства света. Теперь рассмотрим более детально свет с корпускулярной точки зрения. В теории Планка энергия фотона E связана с частотой света ν соотношением

$$E = h\nu. \qquad (1.2)$$

Согласно теории относительности энергия всегда связана с массой соотношением

$$E = mc^2.$$

Поэтому масса фотона

$$m = \frac{h\nu}{c^2}. \qquad (1.3)$$

Поскольку фотоны не существуют в состоянии покоя, то их масса покоя $m_0 = 0$, а масса m, определяемая этой формулой,— это масса фотона, движущегося в вакууме со скоростью c. Импульс фотона равен произведению его массы на скорость:

$$p = mc = \frac{h\nu}{c} = \frac{E}{c}. \qquad (1.4)$$

Отметим, что формула (1.4), дающая связь между энергией фотона E и его импульсом p, является частным случаем общего соотношения, связывающего энергию, импульс и массу покоя любой частицы:

$$\left(\frac{E}{c}\right)^2 = m_0^2c^2 + p^2.$$

Опыт показывает, что свет может взаимодействовать с веществом только путем дискретных процессов, при которых испускается или поглощается целый квант. С корпускулярной точки зрения взаимодействие между веществом и светом описывается как поглощение, испускание или рассеяние фотонов, сопровождающееся изменением их энергии и импульса. Экспериментально рассеяние фотонов на электронах было исследовано Комптоном в 1923 году. В его опытах через вещество с легкими атомами (графит, парафин) пропускался пучок рентгеновских лучей частоты ν. Измерения Комптона показали, что в рассеянном рентгеновском свете, наряду с излучением неизменной длины волны, появляется рентгеновское излучение не-

сколько большей длины волны. Наблюдаемое изменение длины волны $\Delta\lambda = \lambda' - \lambda$ зависит от угла θ между направлением первичного пучка и направлением рассеянного света следующим образом:

$$\Delta\lambda = 2k \sin^2 \frac{\theta}{2}, \qquad (1.5)$$

причем постоянная k, найденная из эксперимента, равна 0,024 Å.

Для объяснения эффекта Комптона применим законы сохранения энергии и импульса к столкновению рентгеновского фотона с электроном. Поскольку в атомах легких элементов энергия связи электрона порядка 10 эВ, что примерно в тысячу раз меньше энергии рентгеновского кванта $h\nu \approx 10$ кэВ, то электроны в этих опытах можно считать практически свободными. Энергия покоя электрона $m_0c^2 = 0,5$ МэВ; поэтому отношение $h\nu/m_0c^2 \ll 1$. Следовательно, покоившийся до столкновения с фотоном или двигавшийся в атоме с нерелятивистской скоростью электрон и после столкновения останется нерелятивистским. Закон сохранения энергии при столкновении в пренебрежении начальной энергией электрона записывается в виде

Рис. 1.1. К объяснению эффекта Комптона.

$$h\nu = h\nu' + \frac{p^2}{2m_0}, \qquad (1.6)$$

где p — импульс электрона после столкновения с фотоном. Закон сохранения импульса

$$\boldsymbol{p}_\Phi = \boldsymbol{p}'_\Phi + \boldsymbol{p}, \qquad (1.7)$$

где $\boldsymbol{p}_\Phi$ и $\boldsymbol{p}'_\Phi$ — импульсы фотона до и после рассеяния, запишем с помощью теоремы косинусов (рис. 1.1) в виде

$$p^2 = p_\Phi^2 + p_\Phi'^2 - 2p_\Phi p'_\Phi \cos\theta. \qquad (1.8)$$

Подставляем в это равенство выражение для импульса фотона через его частоту $p_\Phi = h\nu/c$ и квадрат импульса электрона из закона сохранения энергии (1.6):

$$2m_0 h(\nu - \nu') = \frac{h^2}{c^2}(\nu^2 + \nu'^2 - 2\nu\nu' \cos\theta).$$

Переписав это уравнение в виде

$$\nu - \nu' = \frac{h\nu}{2m_0c^2}\left(\nu + \frac{\nu'^2}{\nu} - 2\nu' \cos\theta\right) \qquad (1.9)$$

и учитывая, что $h\nu/m_0c^2 \ll 1$, видим, что изменение частоты $\Delta\nu = \nu' - \nu$ мало́ по сравнению с самой частотой ν. Поэтому в правой части (1.9) можно ν' заменить на ν. Тогда для относительного сдвига частоты $\Delta\nu/\nu$ при рассеянии сразу получаем

$$\frac{\Delta\nu}{\nu} = -\frac{h\nu}{m_0c^2}(1 - \cos\theta). \qquad (1.10)$$

Знак минус показывает, что частота рентгеновского излучения при рассеянии уменьшается. Это естественно, поскольку фотон отдает часть своей энергии электрону. Теперь перейдем от частот к длинам волн. Это легко сделать, учитывая, что относительное изменение частоты при рассеянии мало: $\Delta\nu/\nu \ll 1$. Тогда

$$\lambda + \Delta\lambda = \frac{c}{\nu + \Delta\nu} \approx \frac{c}{\nu}\left(1 - \frac{\Delta\nu}{\nu}\right).$$

Так как $c/\nu = \lambda$, то из этого выражения имеем $\Delta\lambda = -\lambda\frac{\Delta\nu}{\nu}$, и соотношение (1.10) переписывается в виде

$$\Delta\lambda = \frac{h}{m_0c}(1 - \cos\theta) = \frac{2h}{m_0c}\sin^2\frac{\theta}{2}, \qquad (1.11)$$

Эта формула совпадает с (1.5), ибо величина $h/(m_0c)$, как легко убедиться, как раз равна 0,024 Å. Эта имеющая размерность длины комбинация трех универсальных постоянных получила название комптоновской длины волны электрона.

Интересно отметить, что изменение длины волны рентгеновского излучения в явлении Комптона, как видно из (1.11), не зависит от длины волны падающего излучения. А как объяснить существование в спектре рассеянного излучения еще и несмещенной линии? Все дело в том, что внутренние электроны, особенно в тяжелых атомах, связаны настолько прочно, что их энергия связи уже сравнима с энергией рентгеновских квантов и, следовательно, их нельзя рассматривать как свободные. Поэтому при соуда-

рении фотон обменивается энергией и импульсом с атомом в целом, а так как масса атома очень велика, то по закону сохранения импульса фотон практически не передает ему своей энергии. Следовательно, энергия кванта $h\nu$ при таком рассеянии не изменяется. Гамма-кванты, рассеянные внутренними электронами, образуют несмещенную компоненту, а внешними — смещенную. Из приведенных рассуждений ясно, почему эффект Комптона нельзя наблюдать в видимой области спектра. Вспомните, что энергия фотона видимого света составляет лишь несколько электронвольт. Для такого света даже внешние электроны в легких атомах нельзя считать свободными.

Опыты Комптона ярко продемонстрировали, что энергия и импульс фотона действительно выражаются формулами (1.2) и (1.4), а также то, что законы сохранения энергии и импульса выполняются при индивидуальных процессах рассеяния. Законы сохранения могут быть проверены еще более полно, если исследовать электроны отдачи. Соответствующие измерения показали, что при этом электроны приобретают те же самые импульс и энергию, которые теряет фотон.

Как и в случае фотоэффекта, объяснить особенности явления Комптона, исходя из волновой точки зрения на свет, не удается. Взаимодействуя с классической электромагнитной волной, электрон мог бы получать любое количество энергии. В спектре рассеянного излучения при наблюдении под заданным углом θ можно было бы обнаружить различные значения для изменения длины волны при изменении интенсивности излучения или времени экспозиции. Однако в опытах было однозначно показано, что при заданном угле наблюдается только одно значение смещения длины волны независимо от интенсивности излучения и времени облучения. Эти факты указывают на то, что процесс передачи энергии и импульса не является непрерывным, как это предсказывает классическая теория, а является дискретным, как следует из квантовой теории.

Но не всегда вопрос о природе света стоит так категорично: или кванты, или классические волны. Существует ряд явлений, допускающих корректное объяснение с любой из этих точек зрения. Например, рассмотренный ранее эффект Допплера. Это типично волновое, на первый взгляд, явление может быть объяснено с точки зрения представления о

свете как о потоке фотонов. Покажем это. Пусть «закрепленный» неподвижный атом испускает фотон с энергией $h\nu$ при переходе атома из одного стационарного состояния в другое. Разность энергий этих стационарных состояний $h\nu$ не зависит от того, покоится атом или движется.

Рис. 1.2. К объяснению эффекта Доплера при испускании фотона движущимся атомом.

При испускании фотона свободно движущимся атомом импульс атома изменяется, поскольку испущенный фотон обладает импульсом. Следовательно, кинетическая энергия атома также изменяется. Энергия фотона $h\nu'$, испущенного движущимся атомом, отличается от $h\nu$ вследствие изменения кинетической энергии атома.

На основании закона сохранения энергии

$$h\nu' - h\nu = \frac{p^2}{2M} - \frac{p_1^2}{2M}, \qquad (1.12)$$

где p — импульс атома до испускания фотона, p_1 — после испускания, M — масса атома.

Начальный и конечный импульсы атома можно связать с импульсом испускаемого фотона p_Φ с помощью закона сохранения импульса (рис. 1.2):

$$p = p_1 + p_\Phi. \qquad (1.13)$$

Перенося p_Φ в равенстве (1.13) в левую часть, возводя полученное равенство в квадрат и учитывая, что импульс фотона крайне мал по сравнению с импульсом излучающего атома, получаем

$$p^2 - 2pp_\Phi \cos\theta \approx p_1^2. \qquad (1.14)$$

С помощью (1.14) соотношение (1.12) можно переписать в виде

$$h\nu' - h\nu = \frac{p}{M} p_\Phi \cos\theta. \qquad (1.15)$$

Подставляя в (1.15) импульс испущенного фотона $p_\Phi = h\nu'/c$ и учитывая, что p/M есть скорость движения v излучающего

атома, находим

$$\nu' - \nu = \nu' \frac{v}{c} \cos\theta,$$

откуда

$$\frac{\Delta\nu}{\nu} = \frac{v}{c} \cos\theta \qquad (1.16)$$

с точностью до членов порядка v/c. Это есть обычное, нерелятивистское выражение для явления Допплера.

§ 2. Границы применимости классической физики. Соотношения неопределенностей

Кроме изложенных выше опытных фактов, указывающих на правдоподобность идеи квантования применительно к свету, к началу XX столетия физика накопила и ряд строгих экспериментальных результатов, подтверждающих существование дискретных энергетических уровней у всех атомных систем. В 1913 году Нильс Бор сформулировал правила квантования движения электронов в атомах и с их помощью объяснил известные к тому времени экспериментальные закономерности спектральных линий простейших атомов.

Но в дальнейшем появились еще более ошеломляющие экспериментальные факты. Пучок электронов определенной энергии, прошедший сквозь кристалл, падал на фотопластинку и давал дифракционную картину такую же, как и прошедший сквозь кристалл пучок рентгеновских или γ-лучей. При этом дифракционная картина не зависела от интенсивности электронного пучка; та же картина получалась в предельном случае весьма слабых пучков, когда можно было считать, что электроны падают на кристалл поодиночке. Таким образом, волновые свойства приходилось приписывать каждому электрону в отдельности, а не всей совокупности электронов в пучке. Вместе с тем каждый электрон, попадая на фотопластинку, давал почернение только в одном месте, в одном зерне светочувствительного слоя, и лишь совокупность почерневших зерен давала дифракционную картину распределения интенсивности прошедшего пучка. Итак, в одних условиях при прохождении сквозь кристалл электрон вел себя как

протяженная волна, а в других — при попадании на зерно фотослоя — как строго локализованная частица. На основании представлений классической физики такого различия в поведении электрона в разных условиях объяснить не удалось.

Появились и другие, гораздо менее экзотические на первый взгляд эксперименты, которые тем не менее не допускали объяснения в рамках классической физики. Например, опыты по столкновению атомов водорода с другими частицами показали, что в основном состоянии (а также в некоторых других) атом водорода обладает сферической симметрией — его свойства не зависят от направления. Атом водорода состоит, как известно, из тяжелого ядра — протона и одного электрона. Попробуйте представить себе какую-нибудь конфигурацию из одного положительного и одного отрицательного зарядов, обладающую сферической симметрией!

Все эти явления были правильно объяснены только квантовой механикой. История создания квантовой механики делится на два периода. Первый период — с начала XX столетия и до конца его первой четверти — это период создания так называемой «старой квантовой теории», в основе которой лежат гипотеза Планка о дискретном характере излучения нагретых тел, уравнение Эйнштейна для фотоэффекта и теория атома Бора. Старая квантовая теория не представляла собой стройной, логически замкнутой науки. Удачно описав некоторые экспериментальные факты, она проявила полную неспособность правильно объяснить и количественно описать все многообразие удивительных явлений микромира.

К концу первой четверти XX века стало ясно, что необходима теория, которая с самого начала выяснила бы, почему непригодна классическая физика при анализе явлений атомного масштаба. Какие же черты классического способа описания физических явлений делают его неприменимым к микрообъектам и где проходит граница применимости представлений классической физики?

Классическое описание физического процесса или явления характеризуется рядом абстракций. Прежде всего, это абсолютизация понятия физического процесса, заключающаяся в предположении о независимости явлений от условий их наблюдения. Единственное обстоятельство,

связанное с условиями наблюдения, которое учитывалось в классической физике, есть выбор системы отсчета: по отношению к двум произвольно движущимся друг относительно друга системам отсчета одно и то же явление будет иметь различный вид. Физический процесс в инерциальной системе отсчета рассматривался как нечто, происходящее независимо от наблюдения за этим процессом, а не как явление, конкретно познаваемое при помощи определенных средств исследования. Позднейшее развитие физики показало, что абсолютизация физических процессов не является логически необходимой, а представляет собой допущение, которое прекрасно оправдывалось при изучении макроскопических явлений, но которое оказалось совершенно непригодным в микромире.

Действительно, классическая физика имела дело с телами крупного масштаба; по отношению к которым воздействие, связанное с измерением, играло совершенно ничтожную роль. В тех случаях, когда оно было заметным, его можно было учесть и внести соответствующие поправки. Принципиальная возможность этого никогда не вызывала сомнений.

Вторая абстракция, допускавшаяся в классической физике, была тесно связана с первой и заключалась в том, что при изучении физических явлений считалась возможной сколь угодно подробная детализация описания этих явлений. Другими словами, считалось, что можно неограниченно уточнять наблюдение и наблюдать разные стороны одного и того же физического процесса, не нарушая самого явления.

С этими двумя абстракциями, используемыми в классической физике, с предположением об абсолютном характере физических процессов (в смысле их независимости от условий наблюдения) и о возможности сколь угодно детального их описания (в пределе — исчерпывающе точного и всестороннего), связано понятие о лапласовском механическом детерминизме, согласно которому можно определить состояние исследуемой системы в любой момент времени, коль скоро известно ее начальное состояние.

Вопрос о применимости классического способа описания — это вопрос о возможности использования перечисленных абстракций при анализе конкретного явления. Если в каком-то конкретном случае установлено, что эти

абстракции неприменимы, то классическое описание невозможно, и, следовательно, бессмысленны классические представления о свойствах изучаемого объекта, например о его движении по определенной траектории.

Пределы применимости представлений классической физики, т. е. классического способа описания явлений микромира, устанавливаются так называемыми соотношениями неопределенностей Гейзенберга. Рассматривая различные способы измерения положения и импульса частицы, Гейзенберг пришел к выводу, что условия, благоприятные для точного измерения положения частицы, неблагоприятны для точного измерения ее импульса и, наоборот, условия, благоприятные для измерения импульса, неблагоприятны для измерения положения частицы. Одно из соотношений Гейзенберга связывает между собой неопределенности в значениях координаты частицы x и соответствующей компоненты импульса p_x в один и тот же момент времени:

$$\Delta x \, \Delta p_x \geqslant h. \tag{2.1}$$

Величины Δx и Δp_x неправильно было бы понимать только как неточности одновременного измерения величин x и p_x, поскольку самый термин «неточность» как бы предполагает, что существуют и «точные» значения x и p_x, но только они почему-то не могут быть измерены. На самом деле невозможность точного измерения есть следствие того, что частица по своей природе не имеет одновременно точного значения координаты и соответствующей проекции импульса. Эта невозможность есть проявление корпускулярно-волновой природы материальных микрообъектов. Аналогичные соотношения справедливы и для других координат и компонент импульса:

$$\Delta y \, \Delta p_y \geqslant h, \tag{2.2}$$

$$\Delta z \, \Delta p_z \geqslant h. \tag{2.3}$$

Кроме соотношений (2.1) — (2.3) справедливо соотношение, связывающее неопределенность в изменении энергии частицы и неопределенность в моменте времени, когда это изменение произошло:

$$\Delta E \, \Delta t \geqslant h, \tag{2.4}$$

Это соотношение получило название неравенства Бора — Гейзенберга. Оно фактически означает, что определение энергии с точностью до ΔE должно занять промежуток времени, равный по меньшей мере $\Delta t \sim h/\Delta E$. Таким образом, если изучаемая система находится в некотором состоянии в течение времени Δt, то ее энергия там имеет

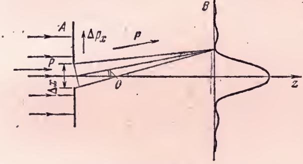


Рис. 2.1. Дифракция света на щели.

неопределенность не менее $\Delta E \sim h/\Delta t$, поскольку Δt — наибольший промежуток времени, в течение которого можно измерять энергию.

Соотношения неопределенностей Гейзенберга являются одним из фундаментальных законов природы. Они справедливы для любых материальных объектов — элементарных частиц, квантов света, атомов, молекул и т. д. Справедливость соотношений неопределенностей, как и всех других фундаментальных законов природы, подтверждается всей совокупностью имеющихся экспериментальных фактов.

Проиллюстрируем неравенство Гейзенберга (2.1), рассматривая дифракцию плоской световой волны на узкой щели.

Пусть на непрозрачный экран A со щелью шириной Δx падает слева плоская монохроматическая волна (рис. 2.1). На удаленном экране B (на рис. 2.1 размер щели сильно преувеличен по сравнению с расстоянием между экранами) наблюдается дифракционная картина, распределение освещенности для которой показано на этом же рисунке.

Почти весь дифрагированный свет приходит в область на экране B, ограниченную главным максимумом. Угловую

ширину этого максимума легко вычислить. Направление на ближайший минимум характеризуется углом θ, определяемым из условия (см. формулу (2.7) на стр. 433)

$$\Delta x \sin\theta = \lambda. \qquad (2.5)$$

Рассмотрим теперь эту дифракционную картину с точки зрения представления о свете как о совокупности световых квантов-фотонов. Каждый фотон, прошедший через щель, попадает в определенную точку на экране B. Предсказать, в какую именно точку попадет отдельный фотон, принципиально невозможно. Однако в совокупности большое число попавших на экран B фотонов дает дифракционную картину, представленную на рис. 2.1. На первый взгляд могло бы показаться, что дифракционную картину можно объяснить интерференцией между различными фотонами, проходящими через щель, т. е. только в рамках корпускулярных представлений. Однако, уменьшая интенсивность света до таких пределов, когда в любой момент времени между источником света и экраном будет находиться в среднем только один фотон, можно убедиться, что распределение фотонов, попавших на экран за достаточно большой промежуток времени, по-прежнему будет определяться дифракционной картиной. Таким образом, дифракция представляет собой статистическое свойство отдельного фотона. Проследим, как происходит движение фотона в этом приборе. До щели в экране A распространяется плоская монохроматическая волна, т. е. нам точно известен импульс фотонов

$$p = \frac{h\nu}{c} = \frac{h}{\lambda}, \qquad (2.6)$$

направленный по оси z. Составляющая импульса фотона по оси x равна нулю, т. е. известна точно, но зато совершенно не определена x-координата фотона. При прохождении фотона через щель в экране A ширина щели Δx будет служить мерой неопределенности значения x-координаты фотона. В самом деле, факт появления фотона на экране B позволяет сделать лишь тот вывод, что фотон проник сквозь щель; в какой же именно точке щели это произошло, совершенно неизвестно. Далее, по корпускулярным представлениям, возникновение на экране дифракционной картины следует истолковать в том смысле, что

каждый фотон, пройдя через щель, отклоняется либо вверх, либо вниз. Но для этого фотон должен приобрести составляющую импульса Δp_x, перпендикулярную направлению первоначального движения. Величина полного импульса фотона p, как видно из формулы (2.6), при этом не меняется, ибо остается неизменной длина волны.

Поскольку бо́льшая часть фотонов попадает в область главного максимума, но принципиально невозможно предсказать, куда попадет каждый фотон, то из рис. 2.1 ясно, что мера неопределенности x-компоненты импульса Δp_x после прохождения через щель есть

$$\Delta p_x \geqslant p \sin\theta. \qquad (2.7)$$

Перемножая почленно (2.5) и (2.7) и учитывая соотношение между импульсом фотона и длиной волны света (2.6), получаем для момента времени, когда фотон проходит через щель:

$$\Delta x\, \Delta p_x \geqslant h,$$

что совпадает с формулой (2.1). Подчеркнем, что проделанный вывод не является доказательством соотношений неопределенностей, а представляет собой лишь иллюстрацию их справедливости для конкретного частного случая.

Соотношения неопределенностей устанавливают принципиальную границу применимости законов классической физики. Используя их, можно выяснить, справедливы ли представления классической физики для описания конкретного явления. Совершенно очевидно, что для макроскопических объектов — планет, искусственных спутников, артиллерийских снарядов — классическое описание является совершенно правильным. Легко убедиться, что при любой достижимой точности измерений координат и импульсов этих объектов соотношения неопределенностей выполняются с огромным запасом и, следовательно, квантовые эффекты никак не проявляются.

Рассмотрим, например, металлический шарик с массой 0,01 г. Если мы определим его положение с точностью $\Delta x \approx 0{,}001$ см, доступной нашему зрению в поле микроскопа, то, согласно соотношению неопределенностей, неопределенность скорости такого шарика равна

$$\Delta v = \frac{\Delta p}{m} \approx \frac{h}{m\, \Delta x} \approx 6 \cdot 10^{-22} \text{ см/с},$$

Такая точность лежит далеко за пределами возможностей измерений.

Посмотрим, как обстоит дело при изучении свойств более мелких объектов, например электронов. Оказывается, что однозначного ответа на вопрос, применимы ли представления классической физики, в этом случае дать нельзя: все зависит от того, какое явление изучается.

Рассмотрим вначале пучок электронов в кинескопе телевизора. В современном телевизоре ускоряющее напряжение $U \approx 15$ кВ. Разогнанный такой разностью потенциалов электрон обладает импульсом $p = \sqrt{2meU}$. Подставляя в эту формулу значения массы электрона m, его заряда e и ускоряющей разности потенциалов, находим в системе СГС: $p = 6{,}6 \cdot 10^{-18}$ г·см/с. Этот импульс направлен вдоль оси трубки. Диаметр пучка, формируемого в современных телевизорах, не бывает меньше $d = 10^{-3}$ см (для телевизора меньший диаметр просто не нужен). Формируя пучок, мы тем самым фиксируем координату электрона в перпендикулярном к оси пучка направлении с точностью Δx, равной диаметру пучка d. В силу соотношения неопределенностей при этом электрону сообщается неконтролируемый импульс Δp, перпендикулярный оси пучка:

$$\Delta p \approx \frac{h}{d} \approx 6{,}6 \cdot 10^{-24} \text{ г·см/с}.$$

Связанная с этим неопределенность в направлении движения электрона $\Delta\theta$ определяется отношением

$$\Delta\theta = \frac{\Delta p}{p} \approx 10^{-6} \text{ рад}.$$

Поскольку длина пути электрона в кинескопе не превышает $l \sim 1$ м, то неконтролируемое смещение ΔS электрона на экране, вызываемое квантовыми эффектами, т. е. неопределённостью в направлении движения электрона $\Delta\theta$, не превосходит $\Delta S \leqslant l\,\Delta\theta = 10^{-4}$ см, т. е. меньше диаметра пучка. Итак, движение электронов в кинескопе телевизора можно рассматривать с помощью законов классической физики.

Рассмотрим теперь электрон в атоме водорода. Хорошо известно, что размер атома водорода d равен приблизительно 10^{-8} см. Классическое описание поведения электрона в атоме предполагает, что ему можно приписать

определенную траекторию. В планетарной модели атома электрон вращается вокруг ядра. Диаметр его орбиты можно считать равным размеру атома. В классической механике условие его движения по круговой орбите радиусом r — второй закон Ньютона — имеет вид

$$\frac{mv^2}{r} = \frac{e^2}{r^2}.$$

Отсюда получаем величину импульса электрона $p = mv$:

$$p = e\sqrt{\frac{m}{r}} \approx 2\cdot 10^{-20}\ \frac{\text{г}\cdot\text{см}}{\text{с}}. \qquad (2.8)$$

Однако с помощью соотношения неопределенностей убеждаемся, что если электрон находится внутри атома, т. е. неопределенность в значении его координаты Δx не превосходит размеров атома d, то соответствующая неопределенность в значении импульса Δp оказывается больше, чем сам импульс, вычисляемый по формуле (2.8):

$$\Delta p \sim \frac{h}{\Delta x} \sim \frac{h}{d} \approx 6{,}6\cdot 10^{-20}\ \frac{\text{г}\cdot\text{см}}{\text{с}}.$$

Итак, для электрона в атоме классическое описание непригодно. Далее на конкретных примерах будет показано, что с помощью соотношений неопределенностей можно не только убеждаться в том, справедливы или нет классические законы в определенных ситуациях, но и исследовать некоторые свойства квантовых объектов.

§ 3. Свет — частицы или волны? Корпускулярно-волновой дуализм. Волны де Бройля

В этом параграфе мы используем соотношения неопределенностей, чтобы разобраться в вопросе: что же все-таки такое свет — частицы или волны? Как мы видели, некоторые оптические явления свидетельствуют в пользу волновых представлений, другие могут быть объяснены только с корпускулярной точки зрения. Наконец, существует целый ряд оптических явлений, которые допускают объяснение как с точки зрения волновых, так и с точки зрения корпускулярных представлений о свете. Рассмотрев несколько примеров, мы могли убедиться, что при анализе конкретных явлений эта двойственность света никак не

мешала нашим рассуждениям и не приводила к логическим противоречиям. Нам только нужно было выбрать, на волновом или корпускулярном языке вести рассмотрение, и последовательно придерживаться выбранного способа описания. Противоречие возникает только тогда, когда мы пытаемся составить общее представление о свете. Действительно, соотношение $E=h\nu$ или $p=h\nu/c$ связывает

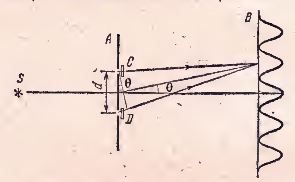


Рис. 3.1. Дифракция света на двух щелях.

волновые и корпускулярные свойства фотона: правые части содержат величину ν, определяемую из интерференционных явлений, а левые части, E и p, характеризуют фотон как частицу. Но именно эти-то свойства света и не могут быть логически непротиворечиво объяснены классической физикой, ибо с точки зрения классической физики понятия волны и частицы являются взаимоисключающими.

Для иллюстрации возникающих логических трудностей, а также для демонстрации того, как они преодолеваются квантовой теорией, рассмотрим подробнее уже упоминавшийся выше простой дифракционный опыт, схематически представленный на рис. 3.1. Источник света S освещает экран A, в котором прорезаны две щели. Расстояние от A до B велико по сравнению с расстоянием d между щелями, которое в свою очередь много больше длины световой волны. На светочувствительном экране B возникает дифракционная картина, причем в местах дифракционных максимумов вырывается наибольшее число фотоэлектронов. Как и в разобранном выше опыте с дифракцией на одной щели, эксперимент показывает, что дифракционная картина

сохранится и в том случае, если в каждый момент времени между источником и экраном в среднем будет находиться только один фотон. Распределение множества фотонов, попавших на экран за достаточно большой промежуток времени, по-прежнему будет определяться классической картиной дифракции от двух щелей, хотя при вырывании фотоэлектронов из экрана B фотоны ведут себя как частицы, каждая из которых выбивает электрон в определенном месте экрана.

Если закрыть одну из щелей, то интерференционные полосы пропадают — распределение интенсивности на экране становится таким же, как при дифракции на одной щели, и при очень узкой щели становится практически равномерным. Поэтому мы вынуждены считать, что при движении от источника света через щели до экрана B излучение ведет себя как волна. Если попытаться объяснить результаты опыта с помощью представления о свете как о частицах, то нужно считать, что каждый фотон, по-видимому, проходит только через одну из щелей. Но тогда, в рамках чисто корпускулярных представлений, можно было бы спросить: каким образом поток независимых фотонов, каждый из которых проходит только через одну из щелей, может образовать дифракционную картину, наблюдаемую лишь при наличии обеих щелей? Или, другими словами, каким образом щель, через которую фотон не проходит, не позволяет ему попасть на те места экрана, куда он мог бы попасть, если бы эта щель была закрыта?

В этой формулировке вопроса предполагается, что фотон действительно проходит через одну из щелей. С точки зрения классической теории это допущение является естественным, ибо предполагается, что в любой момент времени фотон (как и любая другая частица) имеет определенные координаты, доступные измерению. Современная квантовая теория отказывается от этого допущения, утверждая, что говорить о положении фотона имеет смысл лишь в том случае, если при постановке опыта мы позаботимся об определении его координаты. Значит, если мы хотим считать, что каждый фотон действительно, подобно частице, проходит только через одну из щелей, мы должны поставить какой-либо измерительный прибор, который бы фиксировал нам факт прохождения фотона через определенную щель. Если мы попробуем с помощью специальных счетчиков C и D

фиксировать, через какое отверстие проходит каждый фотон, то обнаружим, что дифракционная картина на экране B размоется. Попробуем объяснить этот экспериментальный факт, используя соотношения неопределенностей Гейзенберга. Выясним, можно ли в принципе наблюдать на экране B интерференционные полосы, если точно определять, через какие отверстия проходят фотоны.

Если с помощью счетчиков, установленных непосредственно вблизи отверстий в экране A, мы будем определять, через какое именно отверстие проходит каждый фотон, то тем самым мы действительно заставим свет проявлять корпускулярные свойства, ибо только для частицы имеет смысл утверждение, что она прошла через определенное отверстие, для волны вопрос о том, через какое отверстие она прошла (разумеется, если открыты оба), вообще лишен смысла.

Для того чтобы уверенно судить о том, через какое отверстие прошел фотон, нужно с помощью счетчиков определять его координату x с достаточной точностью, так чтобы ошибка Δx в определении координаты фотона была бы меньше половины расстояния между отверстиями:

$$\Delta x < \frac{d}{2}. \tag{3.1}$$

Однако не следует стремиться определять x-координату фотона слишком точно, так как в силу соотношения неопределенности Гейзенберга

$$\Delta x \, \Delta p_x \gtrsim h \tag{3.2}$$

это приведет к слишком большой неопределенности в значении x-компоненты импульса фотона и дифракционная картина на экране B заведомо размажется, т. е. нам тогда не удастся наблюдать в этом опыте проявление волновых свойств света.

Итак, для наблюдения одновременного проявления светом корпускулярных и волновых свойств необходимо определять координату фотона с максимальной неопределенностью, совместимой с выражением (3.1), т. е. с возможностью установления, через какое отверстие прошел фотон, надеясь при этом, что вносимая при измерении координаты неопределенность в значении импульса Δp_x еще не приведет к полному размытию интерференционных полос. Выясним, совместимы ли эти условия. Какова

максимально допустимая неопределенность в значении импульса Δp_x, которая еще не приводит к полному уничтожению дифракционной картины? Дифракционная картина от двух отверстий, наблюдаемая на экране B, состоит из чередующихся светлых и темных полос. Угловое расстояние $\Delta\theta$ между направлениями на соседние максимум и минимум определяется из условия $d\,\Delta\theta=\lambda/2$, поскольку максимумы расположены в тех точках экрана B, разность хода до которых от отверстий в экране A равна целому числу длин волн. Неопределенность в значении импульса Δp_x можно выразить через неопределенность направления импульса $\Delta\theta_1$: $\Delta p_x=p\,\Delta\theta_1$. Интерференционные полосы, очевидно, не будут полностью размыты, только если $\Delta\theta_1<\Delta\theta$, т. е.

$$\frac{\Delta p_x}{p}<\frac{\lambda}{2d}.$$

Используя соотношение между импульсом фотона и длиной волны $\lambda=h/p$, можно переписать это неравенство в виде

$$2d\,\Delta p_x<h. \qquad (3.3)$$

Таким образом, для проявления волновых свойств света в этом опыте должно выполняться неравенство (3.3), а для проявления корпускулярных свойств — неравенство (3.1). Объединяя эти неравенства, получаем условие одновременного проявления светом корпускулярных и волновых свойств:

$$\Delta x\,\Delta p_x<h/4.$$

Но это условие противоречит соотношению неопределенностей, а потому не может быть выполнено. Итак, установив, через какие отверстия проходят фотоны, мы теряем дифракционную картину и не можем говорить о проявлении фотонами волновых свойств.

Подведем некоторые итоги. Как мы видели, отдельные фотоны обнаруживают волновое поведение, заключающееся в том, что они способны интерферировать сами с собой. Но, попадая на экран, фотоны обнаруживают корпускулярное поведение, заключающееся в том, что они взаимодействуют с веществом только в отдельных точках. Если при этом не делать попытки экспериментально наблюдать траектории фотонов до попадания их на экран, то, пропустив большое число фотонов, мы будем наблюдать на экране

дифракционную картину, предсказываемую волновой теорией. Но предсказать, в какое место экрана попадет определенный фотон, невозможно. Это можно сделать только в вероятностном смысле: вероятность попасть фотону в область максимума освещенности на дифракционной картине велика, а в область минимума — мала. Для проверки такого предсказания нужно большое число фотонов. Если же фиксировать траектории фотонов до попадания на экран, то фотоны вовсе не проявляют волновых свойств.

Как же понимать тот факт, что фотоны появляются иногда в облике частиц, а иногда в облике волн? Квантовая теория отвечает на этот вопрос так: фотон представляет собой квантовый объект, а когда мы описываем его поведение как поведение частицы или волны, мы навязываем классическое описание этому объекту, имеющему существенно неклассическую природу. Рассматривать поведение фотона имеет смысл, только исходя из результатов измерений, совершаемых над ним. Поэтому то, как поведет себя фотон — как частица или как волна, зависит от характера проводимого над ним измерения.

Итак, что же такое свет — частицы или волны? Ни то, ни другое. Мы можем заставить материальный объект, который мы называем светом, проявлять либо корпускулярные, либо волновые свойства. Но в принципе невозможно осуществить эксперимент, в котором свет одновременно проявлял бы и те, и другие свойства. Этим устраняется логическая трудность. Действительно, нам больше не нужно пытаться представить себе, как это фотон может быть одновременно и волной, и частицей. Теперь корпускулярно-волновой дуализм мы понимаем в том смысле, что свет обладает потенциальной возможностью проявлять и волновые, и корпускулярные свойства, но они никогда не проявляются одновременно. Соотношения $E=h\nu$ и $p=h\nu/c$ означают, что между этими взаимно исключающими друг друга свойствами имеет место эквивалентность в том смысле, что меры этих свойств всегда пропорциональны. Эти свойства дополняют друг друга, ибо только их совокупность дает полное представление о свете. Но, как показывают соотношения неопределенностей, в любом явлении в зависимости от конкретных условий реализуется только одна возможность. Корпускулярно-волновой дуализм присущ не только фотонам, но и любым другим микрообъектам.

Мы уже упомннали об экспернментах, в которых иаблюдалась дифракция электроиов при прохождении нх сквозь кристаллы. Но еще до осуществления таких экспернмеитов, в 1924 году, Луи де Бройль предположнл, что все частицы должиы обладать волновыми свойствамн, подобнымн волиовым свойствам света, и ввел колнчественное соотношенне между длнной сопоставляемой частнце волны и импульсом частнцы, аналогичное соотиошеиию между длиной волиы и нмпульсом фотоиа:

$$\lambda = \frac{h}{p}, \tag{3.4}$$

Несколько лет спустя Джермер и Дэвнссон, изучая рассеяние электронов кристаллами, обиаружили дифракцию электронов, подобиую днфракцнн света иа решетке. Атомы крнсталла иикеля, который использовался в опыте, образуют регуляриую конфнгурацню, которая действует подобно дифракцноииой решетке. Максимумы в распределении рассеянных кристаллом электроиов находилнсь в тех местах, для которых выполиялось условие

$$n\lambda = d \sin \theta. \tag{3.5}$$

Но это условне совпадает с нзвестиым условнем для макснмумов прн дифракцнн волн на решетке с пернодом d. Вычнслениое по (3.5) зиачеиие длнны волны λ совпадало со значеннем, даваемым (3.4), с точиостью до 1%. Следовательно, электроны, отражаясь от кристалла, днфрагнруют точио так же, как еслн бы онн были волнамн с длниой, предсказанной де Бройлем.

Результаты этого опыта нмеют фундамеитальиое значенне, ибо оин демоистрнруют волиовые свойства вещества, которые не могут быть пояиты в рамках представленнй о том, что вещество состонт из классическнх частиц. Более позднне опыты показалн, что и другие частнцы вещества, даже такие относительно крупиые, как молекулы, также проявляют волиовые свойства.

Когда реальность проявлеиия волновых свойств частицами стала очевндной, возникла необходимость как-то ннтерпретнровать волны де Бройля, прндать нм определенный фнзическнй смысл. Появнлась концепцня «волны-пнлота», в которой предполагалось, что волна в каком-то смысле «управляет» движением частицы. Но целый ряд

экспериментальных фактов показывал, что такое весьма наглядное представление не приводит к внутренне непротиворечивой картине поведения частиц. Правильное толкование этого вопроса стало возможным только после создания квантовой механики. Пока мы будем просто считать, что соотношение (3.4) дает нам длину волны, которую следует сопоставлять любому материальному объекту, если опыт показывает проявление этим объектом волновых свойств.

Используя эти представления, де Бройль смог объяснить введенное Бором правило квантования электронных орбит в атомах. Для этого он предположил, что допустимые орбиты электронов, например в атоме водорода, соответствуют волне, распространяющейся по кругу около ядра атома. Стационарная волна, представляющая электрон в стационарном состоянии, может быть получена, только если волны непрерывно повторяют себя после каждого полного оборота вокруг ядра. Это условие соответствует, например, условию существования стоячих волн на струне, которые неподвижны и сохраняют свою форму со временем. Для этого на длине орбиты электрона должно укладываться целое число длин волн, т. е.

$$n\lambda = 2\pi r. \qquad (3.6)$$

Подставляя в это соотношение длину волны, выраженную через импульс электрона по формуле (3.4), получим

$$n\frac{h}{p} = 2\pi r, \quad \text{или} \quad mvr = \frac{nh}{2\pi} = n\hbar,$$

что совпадает с правилом квантования орбит электронов в предложенной Бором в 1913 году модели атома водорода.

Представления о волнах де Бройля можно, наряду с соотношениями неопределенностей Гейзенберга, использовать для выяснения вопроса о том, какой теорией, квантовой или классической, следует описывать конкретное явление. Для этого нужно сравнить сопоставляемую по формуле (3.4) изучаемому объекту длину волны с характерными размерами соответствующей задачи, имея в виду, что волновые свойства объекта не играют существенной роли, пока эта длина волны не станет соизмеримой с характерными размерами. Например, длина волны электрона, находящегося на наинизшей боровской орбите в атоме

водорода, в точности равна, как видно из (3.6), расстоянию, проходимому электроном за один оборот вокруг протона. Поэтому квантовые эффекты в этом случае будут весьма существенными и представления классической физики здесь заведомо неприменимы. Сравним теперь длину волны де Бройля, сопоставляемую Земле, с длиной земной орбиты. Так как масса Земли $M=6\cdot10^{27}$ г, скорость Земли на орбите $v\approx30$ км/с, а расстояние от Земли до Солнца $R\approx15\cdot10^{7}$ км, то

$$\frac{\lambda}{2\pi R}=\frac{h}{Mv\cdot2\pi R}\approx3\cdot10^{-75}.$$

Эта величина фантастически мала. Следовательно, движение Земли будет превосходно описываться классической механикой. Любые волновые или квантовые эффекты будут в этом случае меньше, чем, например, эффекты, вызванные столкновением Земли с протоном или электроном, содержащимися в космических лучах.

§ 4. Законы движения в квантовой физике. Принцип соответствия

Итак, мы видели, что многие явления в микромире не описываются классической физикой и, пользуясь соотношениями неопределенностей Гейзенберга, можно установить границы применимости классического способа описания при рассмотрении тех или иных конкретных явлений. В тех случаях, когда классическое описание оказывается непригодным, необходим более совершенный способ описания физических явлений, который должен учитывать возможность проявления изучаемыми объектами как корпускулярных, так и волновых свойств.

Современная квантовая теория ведет свое начало с 1926 года, когда Шрёдингером было предложено уравнение, носящее ныне его имя и лежащее в основе квантовой механики. Разумеется, изложение квантовой механики выходит за рамки этой книги, но мы можем обсудить разобранные выше экспериментальные факты и четко сформулировать, какой должна быть квантовая теория, способная последовательно объяснить все своеобразие явлений микромира.

Обсуждая причины неприменимости представлений классической физики в микромире, мы видели, что эта неприме-

нимость обусловлена рядом абстракций, допускавшихся в классической физике. В классической физике молчаливо предполагалась независимость физических процессов от способов наблюдения и возможность наблюдать одновременно все стороны данного процесса. В области квантовых явлений это не так. Вспомните, например, опыт с дифракцией фотонов на двух щелях: определяя, через какое отверстие проходит каждый фотон, мы этим измерением принципиально изменяли протекание физического процесса, так что дифракционная картина на экране оказывалась полностью размытой.

Анализируя этот опыт, мы приходим к выводу, что основой нового способа описания явлений должен быть явный учет реальных возможностей измерений, проводимых над микрообъектами. Необходимым посредником при изучении таких объектов являются приборы: атомный объект может проявить свои свойства, только провзаимодействовав с прибором. Например, путь микрочастицы становится видимым в результате конденсации паров в камере Вильсона или в результате почернения зерен фотоэмульсии, и т. п. При этом приборы и условия опыта должны описываться классически, путем задания значений параметров, характеризующих приборы. Параметры этих приборов могут, разумеется, задаваться лишь с точностью, допускаемой соотношениями неопределенностей.

В основу нового способа описания поведения микрообъекта следует положить результаты взаимодействия этого объекта с классически описываемым прибором. Свойства атомного объекта выводятся из рассмотрения результатов таких взаимодействий. Это не исключает возможности введения таких величин, которые характеризуют сам микрообъект независимо от прибора (заряд, масса частицы и т. д.), но в то же время позволяет изучать поведение объекта с той его стороны (например, корпускулярной или волновой), проявление которой обусловлено устройством прибора. Таким образом, появляется возможность рассматривать и тот случай, когда разные стороны и разные свойства объекта не проявляются одновременно. По Бору, свойства, проявляющиеся при взаимно исключающих условиях (вспомните обсуждение корпускулярно-волнового дуализма), дополняют друг друга в том смысле, что только их совокупность характеризует объект полностью. Рас-

сматривать одновременное проявление дополнительных свойств не имеет смысла. Таким образом, в новом подходе и не возникает внутреннего противоречия в понятии «корпускулярно-волновой дуализм».

Итак, в основе описания — результаты взаимодействия микрообъекта с прибором. Но опыт показывает, что при данных внешних условиях результат взаимодействия объекта с прибором не является однозначно определенным, а обладает лишь некоторой вероятностью. Вспомните тот же дифракционный опыт: каждый фотон попадал в определенное место экрана, но предсказать точно, в какое именно, было невозможно; существовала лишь определенная вероятность попадания фотона в то или иное место. Таким образом, в описание микрообъекта, его состояния и поведения вводится новый элемент — понятие вероятности, а тем самым и понятие потенциальной возможности.

Понятие вероятности рассматривалось и в классической физике при изучении свойств систем, состоящих из большого числа частиц. Вероятности вводились тогда, когда условия опыта не были полностью известны и по неизвестным параметрам приходилось проводить усреднение. Например, при рассмотрении броуновского движения нам неизвестны координаты и скорости всех молекул, сталкивающихся с броуновской частицей. Поэтому мы могли предсказать только вероятность попадания частицы в то или иное место. В классической физике вероятности отражали неполноту формулировки задачи, которая, быть может, практически и неизбежна, но в принципе устранима. В примере с броуновской частицей нам практически неоткуда взять значения координат и скоростей всех молекул, но классическая физика в принципе допускает возможность измерить все эти величины, не нарушая течения изучаемого процесса, и однозначно предсказать движение броуновской частицы.

В квантовой физике вероятности имеют совсем иной характер. Здесь они принципиально необходимы; их введение характеризует не неполноту условий, а объективно существующие при данных условиях потенциальные возможности. Следует отметить, что в процессе создания квантовой теории высказывалась точка зрения, что введение понятия вероятности и в квантовой механике все-таки связано с тем, что на самом деле микрообъекты обладают

определенной, неизвестной нам внутренней структурой. Поэтому в квантовой механике, как и в статистической физике, приходится считать, что переменные, описывающие эту внутреннюю структуру, распределены случайным образом. Иными словами, в квантовую теорию понятие вероятности вводится только потому, что она не является полной; она станет полной только тогда, когда будут найдены величины, характеризующие внутреннюю структуру частиц. Но в настоящее время можно утверждать, что квантовая теория, основанная на понятии вероятности результата взаимодействия объекта с прибором, оказалась исключительно успешной и, наоборот, нет никаких экспериментальных фактов, которые свидетельствовали бы о неполноте квантовомеханического описания поведения микрообъектов.

Продолжим обсуждение основ квантовой теории. Из сказанного выше ясно, что задать состояние квантового объекта — значит задать распределение вероятностей, или потенциальных возможностей получения того или иного результата взаимодействия объекта с прибором. То обстоятельство, что должны задаваться вероятности также и для тех величин, измерения которых несовместимы, показывает, что речь идет именно о потенциальных возможностях, а не о значениях величин самих по себе, вне связи с условиями их измерения на опыте.

Теперь можно сказать несколько слов и о том, каким должен быть математический аппарат квантовой теории по сравнению с математическим аппаратом классической физики. В классической физике математический аппарат должен давать значения определенных физических величин, т. е. числа. В квантовой теории математический аппарат должен давать не только возможные значения физических величин, но и вероятности получения на опыте тех или иных возможных значений этих величин.

В квантовой теории сформулированные требования удовлетворяются следующим образом. Состояние изучаемой системы характеризуется определенной функцией, называемой волновой или ψ-функцией, в том смысле, что через эту функцию выражаются все вероятности для результатов измерения над системой. Волновая функция определяется из уравнения Шрёдингера — основного уравнения квантовой механики. Каждой физической величине сопостав-

ляются определенные математические операции, которые следует проделать над ψ-функцией, чтобы получить необходимую информацию. Эта совокупность математических операций называется оператором соответствующей физической величины. В квантовой механике каждой физической величине, например энергии, координате, импульсу, сопоставлен определенный оператор. Как выбираются сами операторы? При построении квантовой теории огромную роль сыграл так называемый принцип соответствия, сформулированный Бором: законы квантовой физики должны быть сформулированы таким образом, чтобы в классических границах, когда, например, в изучаемый процесс вовлечено много квантов, эти законы приводили бы к классическим уравнениям для усредненных величин. Использование принципа соответствия позволило найти вид операторов, сопоставляемых определенным физическим величинам. Но общих правил составления операторов для физических величин указать нельзя.

Требование удовлетворения принципу соответствия отнюдь не является тривиальным. Возникает вопрос, как согласовать квантовый подход, основанный на рассмотрении вероятностей, с классическим, допускающим точное предсказание поведения системы. Рассмотрим с этой точки зрения уже разобранный выше пример рассеяния фотонов свободными электронами. Как видно из формулы (1.11), максимально возможное увеличение длины волны при единичном акте рассеяния фотона равно $2h/(m_0c)$, что примерно составляет 10^{-10} см. Это слишком маленькая величина, чтобы ее можно было заметить при рассеянии радиоволн с длиной волны порядка 1 см и больше. Соответствующие таким волнам частоты оказываются порядка 10^{10} Гц, так что энергия одного фотона составляет 10^{-16} эрг, или 10^{-4} эВ. Легко подсчитать, какую наибольшую энергию может приобрести электрон в результате рассеяния одного такого фотона. С помощью формулы (1.10) находим, что $\Delta E_{max}=h\,\Delta\nu_{max}$ составляет 10^{-14} эВ. Для того чтобы в результате рассеяния радиоволн приобрести энергию всего в 1 эВ, электрон должен рассеять по меньшей мере 10^{14} квантов! Разумеется, нельзя точно предсказать результат каждого индивидуального акта рассеяния, но результаты рассеяния такого большого числа квантов, которое фактически представляет собой непрерывный процесс, являются

вполне определенными и совпадают с тем, что дает для этого случая классическая электродинамика.

Проиллюстрируем справедливость принципа соответствия на примере модели Бора для атома водорода. В теории Бора частота света, излучаемого атомом при переходе из стационарного состояния с энергией E_k в состояние с энергией E_n, выражается формулой

$$\nu = \frac{E_k - E_n}{h} = \frac{me^4}{4\pi\hbar^3}\left(\frac{1}{n^2} - \frac{1}{k^2}\right). \qquad (4.1)$$

В применении к атому водорода принцип соответствия означает, что чем больше квантовое число n стационарного состояния, тем лучше выполняются для него законы классической физики. По мере увеличения n радиус орбиты электрона возрастает, а разность энергий двух соседних уровней стремится к нулю. При этом скачкообразные переходы между соседними уровнями становятся почти эквивалентными непрерывному процессу. В этом предельном случае результаты квантовой теории должны совпадать с результатами классической теории.

Применим формулу (4.1) для перехода атома между двумя соседними уровнями E_n и E_{n-1}, считая квантовое число n большим: $n \gg 1$. Тогда

$$\frac{1}{(n-1)^2} - \frac{1}{n^2} = \frac{n^2 - (n-1)^2}{n^2 (n-1)^2} \approx \frac{2}{n^3}$$

и (4.1) для частоты излучаемого света дает

$$\nu = \frac{me^4}{2\pi\hbar^3}\,\frac{1}{n^3}. \qquad (4.2)$$

Используя выражение для радиуса орбиты электрона в n-м состоянии

$$r_n = n^2 \frac{\hbar^2}{me^2},$$

перепишем формулу (4.2) в виде

$$\nu = \frac{e}{2\pi m^{1/2} r_n^{3/2}}. \qquad (4.3)$$

Но точно такое же выражение для частоты света, излучаемого электроном, обращающимся по круговой орбите

радиуса r_n, дает классическая теория. Действительно, с точки зрения классической электродинамики вращающийся по круговой орбите электрон должен излучать электромагнитные волны с частотой ν, равной частоте обращения электрона вокруг ядра. Применяя к движению электрона по круговой орбите с частотой ν второй закон Ньютона,

$$m \cdot 4\pi^2\nu^2 r_n = \frac{e^2}{r_n^2},$$

получаем

$$\nu = \frac{e}{2\pi m^{1/2} r_n^{3/2}}.$$

Итак, в области больших квантовых чисел квантовая теория дает тот же спектр излучения, что и классическая.

Мы построили качественную картину квантовой механики, выяснили на отдельных примерах, что в пределе, при переходе к классическим условиям, результаты квантовой теории переходят в результаты, даваемые классической физикой. С помощью соотношений неопределенностей можно делать оценки границы применимости законов классической физики. Однако эта граница не является четко очерченной, и развитие физики не раз давало примеры вторжения квантовых эффектов в область, кажущуюся абсолютно классической. С некоторыми из этих примеров мы еще встретимся ниже.

АТОМ И ЭЛЕКТРОМАГНИТНОЕ ПОЛЕ

§ 5. Атом в квантовой физике

К находящимся в атомах электронам законы классической физики, вообще говоря, неприменимы. Такие фундаментальные свойства атомов, как их устойчивость, или особенности химического поведения, характеризуемые периодической системой элементов Менделеева, противоречат классической физике и требуют для своего объяснения новых понятий. Поведение электронов в атомах должно описываться квантовомеханически.

Атом с несколькими электронами представляет собой сложную систему взаимодействующих друг с другом электронов, движущихся в кулоновском поле ядра. Согласно

квантовой механике в отсутствие внешних воздействий такая система может находиться только в определенных стационарных состояниях, характеризуемых дискретными значениями энергии. Строго говоря, при этом можно рассматривать только состояния всей системы в целом. Но тем не менее оказывается, что для атома можно приближенно говорить о состояниях каждого электрона в отдельности как о стационарных состояниях электрона в некотором эффективном поле со сферической симметрией, которое создается ядром и всеми остальными электронами. Ясно, что поле, действующее на некоторый электрон, зависит от состояния всех остальных электронов. Поэтому состояния всех электронов в атоме должны определяться одновременно.

Задача определения стационарных состояний электронов в атоме с несколькими электронами не может быть решена точно. Однако существуют различные приближенные методы, которые позволяют с очень высокой точностью рассчитывать наблюдаемые свойства атомов. Единственный атом, для которого эта задача может быть решена точно,— это атом водорода (или водородоподобный ион, т. е. ион с одним электроном, как, например, He^+, Li^{++} и т. д.). Рассчитанные по квантовой механике уровни энергии атома водорода оказались совпадающими с теми, которые давала старая теория Бора. Успех теории Бора связан с тем, что она ввела в физику атома постоянную Планка в виде соотношения $rp \sim h$, связывающего положение и импульс электрона. Все выводы из теории Бора являются следствием этого соотношения, представляющего собой частный случай соотношений неопределенностей Гейзенберга. Многие выводы, по существу, не связаны с используемой в теории Бора классической картиной движения электрона вокруг ядра.

Согласно квантовой механике бессмысленно говорить о движении электрона в атоме по определенной орбите. Физический смысл имеет только вероятность обнаружить электрон в том или ином месте. Квантовомеханическое распределение плотности вероятности местонахождения электрона в атоме можно представить в виде некоторого облака, окружающего ядро атома. Каждому стационарному состоянию, характеризуемому определенным набором квантовых чисел, соответствует свое облако плотности вероятности, имеющее определенную пространственную конфигу-

рацию. Одному и тому же значению главного квантового числа n, определяющего энергию атома $E_n = -me^4/(2\hbar^2 n^2)$, соответствует при $n > 1$ несколько различных состояний с одинаковой энергией, различающихся видом этого облака. Для состояний со сферически симметричным облаком вероятность обнаружить электрон на некотором расстоянии от ядра имеет максимум, когда это расстояние равно радиусу соответствующей боровской орбиты

$$r_n = \frac{\hbar^2}{me^2} n^2.$$

Формулы для энергий стационарных состояний E_n и радиуса электронного облака r_n получены при использовании нерелятивистской механики, условие применимости которой состоит в малости скорости электрона по сравнению со скоростью света. Наибольшей скоростью обладает электрон, находящийся в состоянии с $n=1$, т. е., на языке теории Бора, движущийся по ближайшей к ядру разрешенной орбите. Выразим скорость электрона v через радиус этой орбиты $a_0 = \hbar^2/(me^2)$. Уравнение движения по круговой орбите радиуса a_0 под действием кулоновской силы притяжения к ядру имеет вид

$$\frac{mv^2}{a_0} = \frac{e^2}{a_0^2},$$

откуда, подставляя a_0, находим

$$v^2 = \frac{e^2}{ma_0} = \left(\frac{e^2}{\hbar}\right)^2. \qquad (5.1)$$

С помощью (5.1) находим отношение v/c:

$$\frac{v}{c} = \frac{e^2}{\hbar c} \equiv \alpha. \qquad (5.2)$$

Подставляя сюда численные значения e, $\hbar$ и c, для безразмерной постоянной α получаем значение

$$\alpha = 7{,}3 \cdot 10^{-3} \approx \frac{1}{137}. \qquad (5.3)$$

Видно, что наибольшее возможное значение скорости электрона в атоме водорода в 137 раз меньше скорости света. Таким образом, атом представляет собой нерелятивистскую систему со сравнительно медленно движущимся электроном.

Константа α играет фундаментальную роль в атомной физике. Она известна под названием постоянной тонкой структуры. Такое название объясняется тем, что впервые она появилась в физической теории при нахождении релятивистских поправок к уровням энергии в атоме, которые оказались пропорциональными $(v/c)^2=\alpha^2$.

Постоянная тонкой структуры является одной из истинно фундаментальных констант природы, которая определяет не только релятивистские поправки, но и саму структуру атома. Поясним это. Сравним энергию кулоновского взаимодействия электрона с ядром при расстоянии между ними, равном боровскому радиусу a_0, с энергией покоя электрона mc^2:

$$\frac{e^2}{a_0mc^2}=\frac{e^4}{\hbar^2c^2}=\alpha^2.$$

Мы видим, что характерная для атома энергия связана с энергией покоя электрона через постоянную тонкой структуры. Так как $\alpha^2\ll 1$, то энергия связи электрона в атоме много меньше энергии покоя электрона. Это означает, что атом представляет собой сравнительно «рыхлую», слабо связанную систему.

Взаимодействие между заряженными частицами, пропорциональное e^2, осуществляется через электромагнитное поле. Поэтому постоянная тонкой структуры $\alpha=e^2/\hbar c$ представляет собой безразмерный параметр, характеризующий взаимодействие заряда с электромагнитным полем. Ниже мы увидим, что именно этим параметром определяются особенности излучения электромагнитных волн (света) атомами.

Состояния электронов в более сложных атомах, содержащих несколько электронов, характеризуются таким же набором квантовых чисел, что и у атома водорода. Однако здесь энергия электрона зависит не только от главного квантового числа n, характеризующего размер электронного облака в данном состоянии, но и от других квантовых чисел, характеризующих пространственную конфигурацию этого облака.

Чтобы получить представление обо всей электронной оболочке атома, нужно рассмотреть, каким образом происходит заполнение электронами разрешенных стационарных состояний отдельных электронов. В атоме, находя-

щемся в невозбужденном состоянии, электронные состояния должны быть заполнены таким образом, чтобы энергия всей его электронной оболочки имела наименьшее возможное значение. Конечно, наименьшее значение энергии получилось бы, если бы все электроны находились в наинизшем возможном состоянии, т. е. в состоянии с $n=1$. Но опытные данные свидетельствуют о том, что заполнение электронной оболочки атома происходит иначе.

В 1925 году В. Паули установил общий квантовомеханический принцип, согласно которому в любой системе в каждом разрешенном состоянии не может находиться более одного электрона. Принцип Паули не имеет аналога в классической физике. Он связан с неразличимостью тождественных частиц в микромире: перестановка местами двух электронов в системе не может изменить ее состояния. Принцип Паули дает ключ к пониманию электронной структуры химических элементов и позволяет объяснить тот факт, что химические свойства некоторых атомов с разным числом электронов оказываются сходными. Выражением этого факта является периодичность химических свойств элементов, нашедшая отражение в эмпирически установленной периодической системе Менделеева.

Попытаемся получить общее представление о строении тяжелых атомов, т. е. атомов с большим значением заряда ядра. Электронное облако таких атомов имеет слоистую, или, как говорят, оболочечную, структуру. На первый взгляд могло бы показаться, что размер электронного облака должен монотонно возрастать по мере перехода к все более и более тяжелым атомам, так что тяжелые атомы должны быть гораздо крупнее атома водорода. Однако в действительности это не так. Начнем с «голого» ядра и будем по одному последовательно прибавлять к нему все новые электроны. Для первого электрона, помещенного в поле ядра с зарядом Ze, можно воспользоваться теми же формулами, что и для атома водорода, заменив в них e^2 на Ze^2. Энергия связи такого электрона будет в Z^2 раз больше, чем в атоме водорода, а расстояние от ядра — в Z раз меньше радиуса a_0 первой боровской орбиты в атоме водорода. Это же будет приближенно справедливо и для второго прибавляемого электрона. Но что будет после прибавления n электронов? На больших расстояниях от ядра поле такого иона будет похоже на поле ядра с зарядом

$(Z-n)e$, так как ранее прибавленные электроны, находящиеся близко от ядра, частично экранируют его заряд. Поэтому легко понять, что последующие электроны оказываются всё менее и менее сильно связанными с ядром. После прибавления предпоследнего, т. е. $(Z-1)$-го, электрона электрическое поле иона будет подобно полю облака с суммарным зарядом $+e$. Радиус этого облака будет одного порядка с боровским радиусом a_0. Поэтому энергия связи последнего электрона будет сравнима с энергией связи электрона в атоме водорода, т. е. порядка десяти электрон-вольт, а окончательный размер многоэлектронного атома будет близок к размеру атома водорода и составит несколько ангстрем.

Даже такое грубое представление об электронной структуре тяжелого атома позволяет оценить, к какой области спектра относится излучение, связанное с переходами тех или иных электронов. Легко убедиться, что изменение состояния внешних электронов связано с излучением или поглощением электромагнитных волн оптического диапазона. Поэтому внешние электроны часто называют оптическими. Наоборот, переходы внутренних, близких к ядру атома электронов соответствуют ультрафиолетовой или рентгеновской области спектра.

§ 6. Излучение света атомами. Ширина спектральных линий

Квантовая механика позволяет только на основании известного заряда атомного ядра и числа электронов найти стационарные состояния электронов в свободном атоме и рассчитать многие свойства атомов, в том числе их оптические спектры. Кроме того, на основе квантовой механики можно рассчитать изменение стационарных состояний атома под действием внешних электрических и магнитных полей. Экспериментально эти изменения проявляются в спектрах испускания и поглощения света атомами. Изменения спектра под действием внешнего магнитного поля (явление Зеемана) и под действием электрического поля (явление Штарка) являются эффективными средствами изучения структуры вещества.

Излучаемый атомом свет имеет длину волны, составляющую несколько тысяч ангстрем, что на три порядка пре-

восходит размер атома. Покажем, что большая длина волны излучения является непосредственным следствием малости постоянной тонкой структуры $\alpha=e^2/\hbar c$, характеризующей силу взаимодействия электрона с электромагнитным полем.

Расстояние между уровнями энергии оптического электрона, определяющее согласно формуле $E_2-E_1=\hbar\omega$ частоту излучаемого атомом света, по порядку величины равно энергии связи электрона в атоме водорода $me^4/\hbar^2$. Поэтому длина волны излучения λ равна

$$\lambda=\frac{2\pi c}{\omega}\approx 2\pi\frac{c\hbar^3}{me^4}. \tag{6.1}$$

Выделим в этом выражении боровский радиус $a_0=\hbar^2/me^2$, характеризующий размер излучающего атома. Тогда остающийся множитель $\hbar c/e^2$ равен обратной величине постоянной тонкой структуры:

$$\lambda\approx\frac{2\pi}{\alpha}a_0\approx 1000a_0. \tag{6.2}$$

Большая длина волны излучения по сравнению с размером атома имеет, как мы увидим, важные физические следствия, заключающиеся в том, что спектральные линии излучения имеют малую ширину, а время жизни атома в возбужденном состоянии велико по сравнению с периодом колебаний в излучаемой волне.

Стационарные состояния атомов, определяемые квантовой механикой, являются дискретными. Им соответствуют строго определенные значения энергии. Находящийся в возбужденном стационарном состоянии атом в отсутствие внешних воздействий, согласно квантовой механике, должен оставаться в этом состоянии сколь угодно долго. Однако опыт показывает, что это не так: спустя некоторое время возбужденный атом самопроизвольно переходит в основное состояние, испуская при этом квант света.

Рассмотрим спонтанное излучение атомов подробнее. Покажем, что при спонтанном переходе в основное состояние атом испускает цуг волн конечной протяженности. Это в свою очередь означает, что излучаемый им свет не является строго монохроматическим, а распределен в некотором частотном интервале. Другими словами, спектральные линии, соответствующие спонтанному излучению

атома, имеют некоторую конечную ширину. Это так называемая естественная ширина спектральных линий.

Для простоты будем рассматривать случай, когда спонтанное излучение происходит при переходах электрона между состояниями с большими квантовыми числами. Тогда на основании принципа соответствия процесс излучения света электроном можно рассматривать классически. Как было показано, предсказываемый квантовой теорией спектр излучения атома водорода при больших квантовых числах совпадает со спектром, предсказываемым классической теорией излучения: вращающийся по орбите электрон излучает свет, частота которого равна частоте обращения электрона.

Излучая электромагнитные волны, атом теряет энергию. Радиус орбиты электрона при этом постепенно уменьшается, что на квантовом языке соответствует переходу из состояния с более высокой энергией в более низкое квантовое состояние. Оценим время, в течение которого происходит этот переход.

Движение электрона по круговой орбите можно представить как суперпозицию двух гармонических взаимно перпендикулярных колебаний, сдвинутых друг относительно друга по фазе на $\pi/2$. Амплитуда этих колебаний равна радиусу орбиты, т. е., по порядку величины, размеру излучающего атома. Движущийся с ускорением заряд излучает электромагнитные волны, поэтому с каждым из этих колебаний связана электромагнитная волна. Так как длина волны велика по сравнению с размером атома, т. е. с амплитудой колебаний электрона, то для описания излучения атома можно воспользоваться результатами, полученными при рассмотрении излучения движущегося заряда.

Обе излучаемые вращающимся электроном волны поляризованы во взаимно перпендикулярных направлениях, они не интерферируют, и поэтому их можно рассматривать отдельно. Найдем мощность, уносимую каждой из этих волн. Среднее по времени значение плотности потока энергии волны на расстоянии r от атома дается формулой (14.10) (стр. 412):

$$\langle I\rangle = \frac{1}{32\pi^2\varepsilon_0}\cdot\frac{q^2\omega^4A^2}{c^3r^2}\sin^2\theta.$$

Эта формула написана в системе единиц СИ. Как записать ее в системе единиц СГСЭ, которая используется в этом разделе? Проще всего это сделать следующим образом. Размерность всех величин в этой формуле, кроме заряда q, в обеих системах единиц одинакова, и переход в формуле к системе СГСЭ никак их не затрагивает. Что же касается заряда q, то с этой величиной следует поступить так же, как и в законе Кулона при переходе от единиц СИ к СГСЭ, т. е. заменить $q^2/(4\pi\varepsilon_0)$ на q^2. Так как в рассматриваемом случае величина заряда q равна заряду электрона e, то формула для плотности потока излучаемой энергии принимает вид:

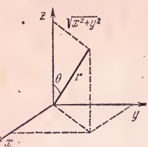


Рис. 6.1. К усреднению по направлениям величины плотности потока излучаемой энергии.

$$\langle j\rangle = \frac{1}{8\pi}\frac{e^2\omega^4A^2}{c^3r^2}\sin^2\theta. \quad (6.3)$$

Чтобы найти полную мощность, излучаемую по всем направлениям, можно усреднить (6.3) по углу θ и умножить полученный результат на $4\pi r^2$. Как видно из рис. 6.1,

$$\sin^2\theta = \frac{x^2+y^2}{r^2}. \quad (6.4)$$

Так как $\langle x^2\rangle = \langle y^2\rangle = \langle z^2\rangle = \frac{1}{3}r^2$, усреднение выражения (6.4) по всем направлениям дает

$$\langle\sin^2\theta\rangle = \frac{2}{3}.$$

Поэтому полная излучаемая осциллятором мощность P равна

$$P = \frac{1}{3}\frac{e^2\omega^4A^2}{c^3}. \quad (6.5)$$

Излучаемая атомом мощность вдвое больше. Заменяя в (6.5) амплитуду колебаний A на боровский радиус a_0 и отбрасывая несущественный при оценках множитель 2/3,

получаем для излучаемой атомом мощности P следующее приближенное выражение:

$$P \approx \frac{e^2\omega^4 a_0^2}{c^3}. \tag{6.6}$$

Теперь мы можем оценить промежуток времени τ, в течение которого продолжается единичный акт излучения, т. е. испускается один фотон с энергией $\hbar\omega$:

$$\tau \approx \frac{\hbar\omega}{P} = \frac{\hbar c^3}{e^2\omega^3 a_0^2}. \tag{6.7}$$

С точки зрения классической физики атом начинает излучать сразу же после того, как он попадает в возбужденное состояние, и этот процесс длится в течение времени τ. По квантовой теории излучение фотона и переход атома в основное состояние происходят мгновенно, скачком. В какой именно момент времени это произойдет, неизвестно. Момент самопроизвольного испускания фотона есть случайная величина, а время τ есть среднее время жизни атома в возбужденном состоянии.

Из формулы (6.7) видно, что обусловленное спонтанным излучением время жизни атома в возбужденном состоянии обратно пропорционально кубу частоты излучения.

Рассмотрим произведение $\omega\tau$, которое с точностью до множителя 2π дает отношение времени жизни возбужденного атома к периоду колебаний. Согласно классическим представлениям $\omega\tau$ определяет число колебаний за время высвечивания. Другими словами, это есть число волн в отдельном излучаемом атомом цуге. С помощью (6.7) имеем

$$\omega\tau \approx \frac{\hbar c^3}{e^2\omega^2 a_0^2}.$$

Произведение ωa_0 равно скорости v электрона на орбите в атоме. Согласно формуле (5.2) отношение v/c равно постоянной тонкой структуры α. Поэтому выражение для $\omega\tau$ принимает вид

$$\omega\tau \approx \frac{\hbar c}{e^2}\,\frac{1}{\alpha^2} = \frac{1}{\alpha^3}. \tag{6.8}$$

Так как $\alpha = 1/137$, то $\omega\tau \approx 10^7$. Для оптических частот $\omega \approx 10^{15}$ с$^{-1}$. Отсюда получаем для времени излучения оценку $\tau \sim 10^{-8}$ с.

Таким образом, в отдельном излучаемом атомом цуге содержится около 10 миллионов волн. Пространственная протяженность такого цуга составляет несколько метров. Это означает, что соответствующая спонтанному излучению свободного атома спектральная линия имеет конечную естественную ширину $\Delta\omega$, связанную с длительностью отдельного цуга τ соотношением (5.5) раздела «Оптика» (стр. 455):

$$\tau\,\Delta\omega \approx 2\pi. \qquad (6.9)$$

Согласно формуле (6.8) отношение естественной ширины $\Delta\omega$ к частоте линии ω равно

$$\frac{\Delta\omega}{\omega} = \frac{\alpha^3}{2\pi} \approx 10^{-7}. \qquad (6.10)$$

На опыте положение уровней энергии атома определяется по наблюдению спектральных линий. Так как линия излучения имеет конечную ширину, то частота излучения ω и, следовательно, энергия возбужденного состояния атома не имеют строго определенного значения. Ширина возбужденного уровня энергии ΔE равна произведению постоянной Планка на ширину спектральной линии $\Delta\omega$, излучаемой при переходе с этого уровня:

$$\Delta E = \hbar\,\Delta\omega. \qquad (6.11)$$

Отношение ширины возбужденного уровня к расстоянию между уровнями, как видно из (6.10), равно всего одной десятимиллионной.

Таким образом, из-за спонтанных переходов возбужденные уровни энергии атома имеют хотя и малую, но конечную ширину, а даваемые квантовой механикой строго дискретные значения энергии этих уровней следует понимать как средние значения. Бесконечно узким является только наинизший уровень энергии, соответствующий основному состоянию атома. Возбужденные состояния атома, строго говоря, не являются стационарными. Они получаются строго дискретными тогда, когда рассматривается изолированный атом. Для получения истинной картины необходимо даже свободный атом рассматривать вместе с электромагнитным полем.

Уширение возбужденных уровней энергии атома вследствие спонтанных переходов находится в полном

соответствии с соотношением неопределенностей Бора—Гейзенберга. Подставляя в (6.11) $\Delta\omega=2\pi/\tau$ из (6.9), получим

$$\Delta E\,\tau \approx 2\pi\hbar = h. \qquad (6.12)$$

Неопределенность в значении энергии возбужденного уровня ΔE связана с неопределенностью в моменте спонтанного перехода в основное состояние τ соотношением (6.12), совпадающим с неравенством Бора — Гейзенберга (2.4) (стр. 532).

Уширение уровней энергии атома проявляется не только в испускании, но и в поглощении света: линии поглощения также имеют конечную ширину. Атом может перейти из основного состояния в возбужденное, поглощая свет в определенном интервале частот, причем ширина этого интервала обусловлена шириной возбужденного уровня энергии.

До сих пор мы рассматривали ширину спектральных линий изолированного атома, находящегося в покое относительно наблюдателя. Эта ширина является внутренним свойством атома. Наблюдаемые на опыте значения ширины спектральных линий, как правило, значительно больше их естественной ширины. Дополнительное уширение линий спектра разреженного газа обусловлено тепловым движением атомов и их столкновениями.

Если среднее время свободного движения атомов между столкновениями τ_1 меньше времени жизни атома в возбужденном состоянии τ, то эффективное время непрерывного излучения цуга волн сокращается до значения τ_1. Ширина линий $\Delta\omega$ в этом случае определяется формулой (6.9), в которой τ заменено на τ_1. Так как частота столкновений зависит от концентрации атомов, то уширение спектральных линий за счет столкновений уменьшается при понижении давления. При низком давлении преобладающим становится уширение, обусловленное явлением Допплера. Из-за теплового движения излучающие атомы имеют различные проекции скорости v_x на направление наблюдения. Поэтому допплеровский сдвиг частоты излучения относительно частоты ω_0 излучения неподвижного атома $(\omega-\omega_0)/\omega_0=v_x/c$ будет различен у разных атомов. В результате наблюдаемая спектральная линия, являющаяся суперпозицией узких естественных линий излучения многих атомов, уширяется. Интенсивность излучения $I(\omega)\,d\omega$ в спектраль-

ном интервале частот от ω до $\omega+d\omega$ пропорциональна числу атомов, излучающих свет в этом интервале. Чтобы найти число таких атомов, нужно в формуле для максвелловской функции распределения по скоростям заменить v_x на $c(\omega-\omega_0)/\omega_0$. В результате получим

$$I(\omega)\,d\omega = I_0 \exp\left\{-\frac{Mc^2}{2kT}\left(\frac{\omega-\omega_0}{\omega_0}\right)^2\right\}d\omega, \tag{6.13}$$

где I_0 — интенсивность излучения на частоте ω_0, а M — масса атома. Контур спектральной линии, описываемый формулой (6.13), показан на рис. 6.2. Принимая за ширину линии интервал частот $\Delta\omega$, на границах которого интенсивность в e раз меньше, чем в центре линии, получим

Рис. 6.2. Контур спектральной линии, обусловленный допплеровским уширением.

$$\frac{Mc^2}{2kT}\left(\frac{\Delta\omega}{2\omega_0}\right)^2 = 1,$$

откуда для допплеровской ширины линии имеем

$$\Delta\omega = \frac{2\omega_0}{c}\sqrt{\frac{2kT}{M}} = 2\omega_0\frac{\sqrt{\langle v^2\rangle}}{c}. \tag{6.14}$$

Допплеровская ширина спектральных линий пропорциональна корню из абсолютной температуры газа. Так как допплеровская ширина значительно превышает естественную, то увидеть естественную ширину линии можно лишь в специальных условиях, например наблюдая излучение пучка атомов в направлении, перпендикулярном их скорости.

§ 7. Излучение света нагретыми телами

Спектральный состав излучения отдельных возбужденных атомов представляет собой набор сравнительно узких линий. Это значит, что излучаемый разреженными газами или парами свет концентрируется в узких спектральных интервалах вблизи определенных частот, характерных для атомов каждого сорта.

Совсем иной вид имеет спектр излучения твердых и жидких тел, нагретых до высокой температуры. В этом излучении, называемом тепловым, присутствуют электромагнитные волны всех частот из очень широкого диапазона, т. е. его спектр является сплошным.

Чтобы получить представление о характере теплового излучения, рассмотрим несколько тел, нагретых до различной температуры, помещенных в замкнутую полость, внутренние стенки которой полностью отражают падающее на них излучение. Опыт показывает, что такая система рано или поздно приходит в состояние теплового равновесия, при котором все тела приобретают одинаковую температуру. Так происходит и в том случае, если внутри полости будет абсолютный вакуум и тела могут обмениваться энергией только путем излучения и поглощения электромагнитных волн. Это позволяет применить при изучении такой системы законы термодинамики. В равновесии все тела в единицу времени поглощают столько же энергии электромагнитных волн, сколько излучают, а плотность энергии излучения, заполняющего полость, достигает некоторой определенной величины, соответствующей установившейся температуре. Такое излучение, находящееся в термодинамическом равновесии с телами, имеющими определенную температуру, называется равновесным или черным излучением. Не только плотность энергии, т. е. полная энергия единицы объема, но и спектральный состав равновесного излучения, заполняющего полость, зависит только от температуры и совершенно не зависит от свойств тел, находящихся в полости.

Для экспериментального изучения спектрального состава равновесного излучения можно проделать небольшое отверстие в окружающей полость оболочке. Излучение, выходящее наружу через отверстие, хотя и не является равновесным, обладает тем не менее в точности таким же спектральным составом, что и заполняющее полость равновесное излучение. Выходящее из отверстия излучение отличается от равновесного только тем, что оно не является изотропным, так как распространяется в определенном направлении.

Если увеличивать температуру в полости, то будет возрастать уносимая выходящим из отверстия излучением энергия. Это означает, что объемная плотность энергии

равновесного излучения растет с температурой. Этот рост происходит очень быстро, как мы увидим ниже, пропорционально четвертой степени абсолютной температуры. С увеличением температуры изменяется и спектральный состав излучения, причем таким образом, что максимум смещается в область более коротких волн: выходящий из отверстия в горячей печи свет имеет красноватый оттенок при сравнительно невысокой температуре и становится желтым и даже белым по мере ее роста.

Что можно увидеть, заглянув через отверстие внутрь полости, в которой излучение находится в равновесии с телами? Так как свойства выходящего из отверстия излучения при тепловом равновесии не зависят от природы находящихся внутри полости тел, то излучение не может нести никакой информации об этих телах, кроме их температуры. И действительно, заглянув внутрь печи, мы не увидим ни предметов на фоне стенок полости, ни самих стенок, хотя в глаз будет попадать много света. Контуры предметов внутри полости не будут видны, все будет представляться одинаково светлым.

Возможность различать предметы появляется только при использовании неравновесного излучения. Если даже это излучение исходит от раскаленных тел и его спектральный состав близок к равновесному, температура излучающей поверхности должна быть выше температуры освещаемых предметов.

Равновесное тепловое излучение можно рассматривать как газ, состоящий из фотонов. Фотонный газ является идеальным, так как разные электромагнитные волны в вакууме при распространении не взаимодействуют друг с другом. Поэтому установление теплового равновесия в фотонном газе возможно только при наличии материальных тел. Механизм установления теплового равновесия заключается в поглощении одних и испускании других фотонов веществом. Возможность поглощения и испускания фотонов приводит к характерной особенности фотонного газа: число частиц в нем не является постоянным, а само определяется из условия термодинамического равновесия.

Представление о фотонном газе позволяет очень просто найти зависимость плотности энергии равновесного излучения w от абсолютной температуры T. Это можно сделать, воспользовавшись соображениями размерности. Энергию

единицы объема излучения можно представить в виде произведения среднего числа фотонов в единице объема n, равномерно заполняющих полость, на среднюю энергию одного фотона $\langle E \rangle$:

$$w = n\langle E \rangle. \tag{7.1}$$

Величины, от которых может зависеть средняя энергия фотона и число фотонов в единице объема равновесного излучения,— это абсолютная температура T, постоянная Больцмана k, скорость света c и постоянная Планка $\hbar$. Поскольку равновесное излучение в полости не зависит ни от размеров и формы полости, ни от природы тел, находящихся в полости, и вещества ее стенок, то такие параметры, как размеры тел и полости, и такие константы, как заряды и массы электронов и ядер, не могут фигурировать в выражениях для n и $\langle E \rangle$.

Средняя энергия фотона теплового излучения по порядку величины равна kT: $\langle E \rangle \approx kT$. Размерность числа фотонов в единице объема n есть L^{-3}. Из величин T, k, c и $\hbar$ можно составить единственную комбинацию, имеющую размерность длины: это $c\hbar/kT$. Поэтому концентрация фотонов n пропорциональна величине $(kT/\hbar c)^3$. Подставляя это выражение в (7.1), можем написать

$$w = C_1 \frac{k^4}{\hbar^3 c^3} T^4, \tag{7.2}$$

где C_1 — некоторый безразмерный множитель.

Формула (7.2) показывает, что объемная плотность энергии равновесного излучения пропорциональна четвертой степени абсолютной температуры в полости. Такой быстрый рост плотности энергии с температурой обусловлен не столько ростом средней энергии фотонов (которая пропорциональна T), сколько увеличением числа фотонов в полости, которое пропорционально кубу температуры.

Если в стенке полости имеется небольшое отверстие, то поток энергии излучения j через единицу площади отверстия пропорционален произведению плотности энергии в полости на скорость света c:

$$j = C_2 \frac{k^4}{\hbar^3 c^2} T^4 = \sigma T^4, \tag{7.3}$$

где σ носит название постоянной Стефана — Больцмана. Точный расчет, основанный на применении статистиче-

ской механики к фотонному газу, дает для нее значение, равное

$$\sigma = \frac{\pi^2 k^4}{60\hbar^3 c^2} = 5{,}67 \cdot 10^{-5}\ \text{г/(с}^3 \cdot \text{К}^4\text{)}.$$

Таким образом, полная интенсивность излучения из отверстия пропорциональна четвертой степени абсолютной температуры в полости.

Излучение с поверхности нагретых тел отличается от излучения из отверстия в стенке полости. Интенсивность и спектральный состав этого излучения зависят не только

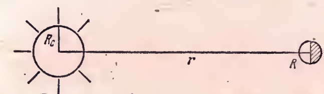


Рис. 7.1. К оценке средней температуры земной поверхности.

от температуры, но и от свойств излучающего тела. Но во многих случаях при оценках можно считать, что эти отличия невелики.

В качестве примера применения закона теплового излучения (7.3) рассмотрим вопрос о средней температуре земной поверхности. Будем считать, что тепловой баланс Земли определяется главным образом поглощением энергии солнечного излучения и излучением энергии в пространство, а роль процессов, идущих внутри Земли, невелика. Полный поток энергии, излучаемой Солнцем, в соответствии с (7.3) равен $\sigma T_C^4 \cdot 4\pi R_C^2$, где T_C — температура поверхности Солнца, а R_C — его радиус. Будем считать, что вся энергия солнечного излучения, падающая на Землю, поглощается. С помощью рис. 7.1 легко сообразить, что количество поглощаемой Землей в единицу времени энергии равно

$$\sigma T_C^4 \cdot 4\pi R_C^2 \cdot \frac{\pi R^2}{4\pi r^2}. \tag{7.4}$$

Такое же количество энергии Земля должна излучать, иначе ее температура не будет оставаться постоянной. Но излучаемая с поверхности Земли энергия также находится

с помощью формулы (7.3), в которую в этом случае следует подставить температуру земной поверхности T. Приравнивая поток излучаемой Землей энергии $\sigma T^4 \cdot 4\pi R^2$ поглощаемой ею энергии (7.4), найдем

$$T = T_С \cdot \frac{1}{2} \sqrt{\frac{2R_С}{r}}. \qquad (7.5)$$

Из этого выражения видно, что для определения T необходимо знать только температуру поверхности Солнца и угловой диаметр Солнца $2R_С/r$, видимый с Земли. Этот диаметр равен 0,01 радиана, а температура поверхности Солнца составляет примерно 6000 К. По формуле (7.5) находим $T \approx 300$ К.

В заключение отметим, что спектр излучения нагретых тел является настолько широким, что коэффициент полезного действия ламп накаливания и других осветительных приборов, основанных на излучении раскаленных тел, совершенно ничтожен. Область видимого света соответствует лишь узкой полосе в спектре теплового излучения.

§ 8. Вынужденное излучение. Квантовые усилители и генераторы света

Спонтанное излучение возбужденного атома с точки зрения квантовой теории происходит скачком. Предсказать момент испускания атомом фотона невозможно, можно говорить только о среднем времени жизни атома в возбужденном состоянии τ.

Пусть теперь мы имеем совокупность очень большого числа атомов, которые образуют настолько сильно разреженный газ, что взаимодействием между атомами можно пренебречь. Каждый атом может находиться в различных квантовых состояниях, которым соответствует дискретный ряд уровней энергии. Будем для простоты считать, что атом может находиться только в двух состояниях: основном с энергией E_1 и возбужденном с энергией E_2. Если какой-нибудь атом в момент времени t находится в возбужденном состоянии, то в следующий интервал времени Δt он может либо остаться в этом состоянии, либо перейти скачком в основное состояние, испустив фотон с энергией $E_2 - E_1$. Обозначим вероятность спонтанного перехода атома в

единицу времени из возбужденного состояния в основное через A_{21}.

Пусть в момент времени t в возбужденном состоянии находятся N_2 атомов. Так как атомы не взаимодействуют между собой, то переход в основное состояние они совершают независимо друг от друга. Поэтому число испущенных за промежуток времени Δt фотонов Δn пропорционально N_2:

$$\Delta n = N_2 A_{21} \Delta t. \qquad (8.1)$$

Если при этом никаких процессов возбуждения атомов не происходит, то изменение числа возбужденных атомов ΔN_2, согласно (8.1), равно

$$\Delta N_2 = -N_2 A_{21} \Delta t. \qquad (8.2)$$

Вероятность спонтанного перехода в единицу времени A_{21} не зависит от времени. Поэтому решение уравнения (8.2) $N_2'(t) = -N_2(t) A_{21}$ имеет вид

$$N_2(t) = N_{20} e^{-A_{21} t}, \qquad (8.3)$$

где N_{20} — некоторая постоянная. Полагая в (8.3) $t=0$, видим, что N_{20} равно числу возбужденных атомов при $t=0$.

Таким образом, в отсутствие внешних воздействий число возбужденных атомов убывает со временем по экспоненциальному закону. Промежуток времени τ, в течение которого число возбужденных атомов в результате спонтанного излучения уменьшается в e раз,— это есть среднее время жизни атома в возбужденном состоянии. Как видно из формулы (8.3),

$$\tau = \frac{1}{A_{21}}. \qquad (8.4)$$

Вероятность спонтанного излучения обратно пропорциональна времени жизни τ. Как следует из формулы (6.7), эта вероятность пропорциональна кубу частоты излучения.

Независимый, случайный характер процессов спонтанного излучения атомов проявляется в том, что различные атомы излучают неодновременно и независимо. Поэтому фазы, направления распространения и состояния поляризации разных цугов волн не согласованы друг с другом.

Это приводит к очень важному следствию: спонтанное излучение некогерентно.

В присутствии электромагнитного поля, кроме спонтанного излучения, будут происходить и процессы возбуждения атомов, т. е. переходы из основного состояния в возбужденное с поглощением фотонов с энергией $\hbar\omega=E_2-E_1$. Очевидно, что вероятность такого перехода пропорциональна объемной плотности энергии электромагнитного поля $w(\omega)$ на частоте перехода и некоторому коэффициенту B_{12}, характеризующему вероятность возбуждения данного атома. Полное число переходов Δn из основного состояния в возбужденное за промежуток времени Δt пропорционально числу атомов в основном состоянии N_1:

$$\Delta n=N_1B_{12}w(\omega)\Delta t. \tag{8.5}$$

Если атомы находятся в термодинамическом равновесии с электромагнитным полем, то число переходов с поглощением фотонов и с их испусканием должно быть одинаково. Основываясь на этом, можно, приравнивая (8.1) и (8.5), найти вид функции $w(\omega)$, т. е. спектральное распределение энергии равновесного излучения. Однако получающийся при этом результат не согласуется с экспериментом и с формулой Планка. Эйнштейн впервые показал, что для получения согласующегося с опытом результата необходимо предположить, что электромагнитное поле вызывает не только переходы атомов из основного состояния в возбужденное, но и переходы из возбужденного состояния в основное с излучением фотонов. Такие переходы, в отличие от спонтанных, получили название индуцированного или вынужденного излучения. Число вынужденных переходов за время Δt пропорционально плотности энергии электромагнитного поля на частоте перехода, числу атомов в возбужденном состоянии N_2 и некоторому коэффициенту B_{21}, характеризующему вероятность такого перехода в атоме. С учетом спонтанного излучения полное число переходов за время Δt из возбужденного состояния в основное равно

$$\Delta n=[N_2A_{21}+N_2B_{21}w(\omega)]\Delta t. \tag{8.6}$$

В состоянии термодинамического равновесия следует приравнять правые части в выражениях (8.5) и (8.6) и учесть, что отношение числа атомов в возбужденном состоянии к

числу атомов в основном состоянии, в соответствии с равновесным статистическим распределением Гиббса (§ 3 раздела «Молекулярная физика», стр. 147), равно

$$\frac{N_2}{N_1} = \exp\left(-\frac{E_2 - E_1}{kT}\right) = \exp\left(-\frac{\hbar\omega}{kT}\right). \qquad (8.7)$$

Легко убедиться в том, что вероятности B_{12} и B_{21} для атома данного сорта равны друг другу. Действительно, при очень высокой температуре, когда $kT \gg \hbar\omega$, плотность энергии w, пропорциональная четвертой степени температуры, становится настолько большой, что в формуле (8.6) можно пренебречь первым слагаемым по сравнению со вторым. Это означает, что в равновесии при высокой температуре вынужденное излучение преобладает над спонтанным. Приравнивая для этих условий правые части (8.5) и (8.6), имеем

$$N_1 B_{12} = N_2 B_{21}. \qquad (8.8)$$

Так как в равновесии при $kT/\hbar\omega \to \infty$, согласно (8.7), $N_1 \to N_2$, то из (8.8) получаем $B_{12} = B_{21}$. Вероятности B_{12} и B_{21} зависят только от свойств атома и не зависят от внешних условий, в которых происходят переходы. Поэтому полученное при $T \to \infty$ равенство B_{12} и B_{21} справедливо всегда, в том числе и в отсутствие теплового равновесия.

В последнее время вынужденные переходы нашли важное практическое применение. На их основе созданы квантовые генераторы излучения — мазеры, генерирующие в микроволновом диапазоне, и лазеры, излучающие в оптическом диапазоне от инфракрасных до ультрафиолетовых лучей. Наиболее существенная особенность квантовых генераторов, с которой связаны практически все их замечательные свойства,— это когерентность создаваемого ими излучения.

Дело в том, что волны, испущенные в результате вынужденных переходов, обладают следующей важной особенностью: их частота, фаза, направление распространения и состояние поляризации точно такие же, как и у излучения, вызвавшего переходы. Иначе говоря, фотоны, испущенные при вынужденном излучении, неотличимы от фотонов, вызывающих это излучение. Поэтому при индуцированном излучении увеличиваются только амплитуда и энергия волны.

Пояснить когерентный характер вынужденного излучения можно следующим образом. Процесс вынужденного излучения является обратным по отношению к поглощению. При распространении пучка лучей в поглощающей среде происходит уменьшение интенсивности, но полностью сохраняются свойства когерентности. Это видно хотя бы из того, что в любых интерференционных опытах прохождение пучка света через серый фильтр, уменьшающий интенсивность без изменения спектрального состава, не разрушает интерференционной картины. Поэтому можно ожидать, что при прохождении света через среду, содержащую возбужденные атомы, в результате вынужденного излучения будет происходить усиление распространяющейся волны при сохранении ее когерентности.

Пусть параллельный пучок монохроматического излучения частоты ω, соответствующей разности каких-либо двух уровней энергии атомов среды, распространяется сквозь эту среду. Изменение числа фотонов в потоке на протяжении расстояния $\Delta x = c\,\Delta t$ за счет процессов поглощения и вынужденного излучения равно

$$\Delta n = [N_2 B_{21} w(\omega) - N_1 B_{12} w(\omega)] \frac{\Delta x}{c}. \qquad (8.9)$$

Вкладом спонтанного излучения в рассматриваемый поток фотонов можно пренебречь, так как спонтанное излучение распространяется по всем направлениям и в направлении рассматриваемого пучка окажется ничтожная его часть. Так как $B_{12} = B_{21}$, то (8.9) можно переписать в виде

$$\Delta n = (N_2 - N_1)\, B_{12}\, w(\omega) \frac{\Delta x}{c}. \qquad (8.10)$$

Эта формула позволяет понять основную идею работы квантового усилителя. Если число атомов в возбужденном состоянии N_2 меньше числа атомов в основном состоянии N_1, то при распространении волны поглощение будет преобладать над вынужденным излучением и интенсивность волны будет убывать. Так обычно и обстоит дело, если волна распространяется в среде, находившейся в состоянии термодинамического равновесия или близком к нему.

Но если в силу каких-то причин число возбужденных атомов N_2 превосходит число атомов в основном состоя-

нии N_1, то по мере распространения волны число фотонов в пучке будет нарастать — волна будет усиливаться.

Найдем закон, по которому будет изменяться интенсивность волны по мере ее распространения в среде. Произведение объемной плотности энергии волны $w(\omega)$ на скорость света c есть плотность потока энергии, переносимой волной: $j=cw(\omega)$. Так как изменение потока фотонов на протяжении отрезка Δx равно $c\,\Delta n$, то изменение потока энергии Δj на этом же отрезке Δx равно $\hbar\omega c\,\Delta n$. Умножая обе части равенства (8.10) на $\hbar\omega c$, найдем

$$\Delta j = \frac{\hbar\omega}{c}(N_2 - N_1)\,B_{12} j\,\Delta x. \qquad (8.11)$$

Переходя к пределу при $\Delta x \to 0$, получим уравнение для функции $j(x)$ — плотности потока энергии — в виде

$$\frac{dj}{dx} = \alpha j(x), \qquad (8.12)$$

где не зависящий от x коэффициент α равен

$$\alpha = \frac{\hbar\omega}{c}(N_2 - N_1)B_{12}. \qquad (8.13)$$

Решение уравнения (8.12) имеет вид

$$j(x) = j_0 e^{\alpha x}. \qquad (8.14)$$

Здесь j_0 есть плотность потока энергии пучка при $x=0$. При $\alpha>0$ поток энергии нарастает по мере распространения волны, а при $\alpha<0$ убывает. Если $N_2>N_1$, то $\alpha>0$. В этом случае среда называется активной.

Каким же образом удается получить сильно неравновесное состояние вещества, в котором $N_2>N_1$? В зависимости от типа рабочего вещества этого можно добиться разными методами. В импульсных твердотельных квантовых генераторах, например в рубиновом лазере, используется оптическая накачка светом мощной импульсной лампы-вспышки. В полупроводниковых лазерах непрерывного действия неравновесное состояние достигается при пропускании электрического тока через $p-n$-переход. В газовых лазерах атомы рабочего вещества возбуждаются электрическим разрядом.

В квантовых генераторах для получения монохроматического когерентного излучения, помимо эффекта усиления волны при прохождении через активную среду, используется положительная обратная связь. Часть излучаемой световой энергии все время должна оставаться внутри рабочего вещества, вызывая вынужденное излучение света все новыми и новыми возбужденными атомами. Это осуществляется с помощью зеркал. Обычно рабочее вещество помещается между двумя параллельными зеркалами, одно из которых полупрозрачно (рис. 8.1). Такую систему называют оптическим резонатором.

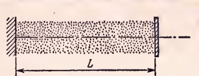


Рис. 8.1. Оптический резонатор с активной средой.

Возникшая в каком-либо месте в результате спонтанного перехода световая волна усиливается за счет вынужденного испускания при распространении через активную среду. Эффективно будут усиливаться только те волны, направление распространения которых совпадает с осью резонатора, так как при всех других направлениях волна быстро покидает пределы активной среды. Дойдя до полупрозрачного зеркала, волна частично выходит наружу, а частично отражается назад. Отразившаяся волна дает начало новой лавине фотонов. Пройдя вдоль резонатора от одного зеркала до другого путь L через активную среду, волна, в соответствии с формулой (8.14), усиливается в $e^{\alpha L}$ раз. Отразившаяся от второго зеркала волна снова на длине резонатора усиливается в $e^{\alpha L}$ раз и т. д. Таким образом, наличие зеркал увеличивает эффективное расстояние, которое распространяющаяся вдоль оси резонатора волна проходит в активной среде.

Очевидно, что генерация возможна только тогда, когда падающая на полупрозрачное зеркало после очередного прохода волна имеет энергию не меньшую, чем при предыдущем падении. Это значит, что усиление света в активной среде должно быть достаточно большим, превышающим некоторое значение, называемое пороговым.

Оптический резонатор, образованный зеркалами, помимо осуществления необходимой для возникновения генерации положительной обратной связи, выполняет и другую

важную функцию — формирует когерентное монохроматическое излучение. Для того чтобы выходящее через полупрозрачное зеркало излучение было когерентным, необходимо, чтобы составляющие его последовательные волновые цуги были согласованы друг с другом. К полупрозрачному зеркалу подходят тождественные цуги волн, возникшие при вынужденных переходах в активной среде из единственного спонтанно испущенного цуга. Часть из них выходит наружу, часть отражается. Отразившиеся цуги проходят через резонатор туда и обратно и снова возвращаются к полупрозрачному зеркалу. Образовавшаяся в результате разность хода должна быть равна целому числу длин волн:

$$2L=n\lambda, \quad n=1, 2, \ldots \tag{8.15}$$

Только в этом случае выходящее наружу излучение будет строго монохроматическим. Отсюда следует, что квантовый генератор может создавать монохроматическую волну не произвольной частоты, а лишь с дискретным набором частот ω_n, соответствующих допустимым значениям длины волны $\lambda_n=2L/n$. Этот набор частот определяется формулой

$$\omega_n=\frac{2\pi c}{\lambda_n}=\frac{\pi c}{L}n. \tag{8.16}$$

Оптический резонатор, образованный зеркалами, можно рассматривать как колебательную систему, в которой собственные нормальные колебания (моды) имеют вид стоячих электромагнитных волн с узлами на зеркалах. Частоты таких колебаний определяются точно таким же условием (8.16). Поэтому квантовый генератор можно рассматривать как автоколебательную систему, в которой возможны незатухающие колебания на одной из собственных частот резонатора.

Возбужденный уровень энергии атомов активной среды всегда имеет конечную ширину. Поэтому усиление света при вынужденных переходах происходит не на одной фиксированной частоте, а в некотором интервале частот $\Delta\omega$, определяемом шириной спектральной линии. Обычно на интервал $\Delta\omega$ приходится несколько собственных частот резонатора (рис. 8.2). Некоторые из них могут оказаться

в интервале $\Delta\omega_1$, в котором величина коэффициента усиления превосходит пороговое значение. На этих частотах и возможна генерация излучения. Если в интервал $\Delta\omega_1$ попадает только одна собственная частота резонатора, то, казалось бы, лазер непрерывного действия должен генерировать строго монохроматический свет. Однако по ряду

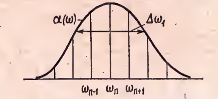


Рис. 8.2. В пределах контура спектральной линии находится несколько собственных частот резонатора.

причин это, конечно, не так, хотя излучение такого одномодового лазера обладает очень высокой степенью монохроматичности. И действительно, ширина линии лазерного излучения может быть много меньше не только доплеровской, но и естественной ширины спектральных линий. Ширина лазерной линии может быть доведена до значения 10^{-14} от самой частоты. Соответствующая такой монохроматичности длина цуга когерентных волн достигает сотен километров.

Наиболее существенной причиной нестабильности частоты генерации являются случайные изменения длины резонатора L, вызываемые, например, тепловым расширением, вибрациями и т. п.

Замечательной чертой лазеров, тесно связанной с когерентностью создаваемого ими излучения, является исключительная способность к концентрации световой энергии: к концентрации в спектре — очень узкая спектральная линия излучения, концентрации во времени — возможность получать сверхкороткие импульсы света, к концентрации в пространстве и по направлению распространения — возможность получить практически параллельный пучок, расходимость которого определяется только дифракционными эффектами и поэтому очень мала, и сфокусировать все излучение в малой области с размерами порядка длины волны.

КВАНТОВАЯ ФИЗИКА И СВОЙСТВА МАКРОСКОПИЧЕСКИХ ТЕЛ

§ 9. Электронная структура кристаллов. Диэлектрики, полупроводники, металлы

Любое макроскопическое тело состоит из атомов и молекул, подчиняющихся законам квантовой физики. Поэтому объяснение наблюдаемых свойств макроскопических тел на основе представлений об их микроструктуре невозможно без использования квантовых законов.

Наибольший прогресс в применении квантовой механики к макроскопическим системам достигнут при изучении твердых тел, т. е. тел, обладающих кристаллической структурой. Первое, что нужно знать для объяснения наблюдаемых электрических, магнитных, тепловых, оптических и других макроскопических свойств,— это уровни энергии электронов в кристалле. Строго говоря, в кристалле, как и в отдельном атоме, можно рассматривать только состояния всей системы в целом. Тем не менее, как и в атоме, с хорошей точностью можно говорить о состояниях отдельных электронов в некотором эффективном пространственно периодическом поле кристалла.

Получить качественное представление о структуре энергетического спектра электронов в твердом теле можно, проследив за тем, как уровни энергии изолированных атомов изменяются при объединении этих атомов в кристалл. Допустим, что N одинаковых атомов расположены в пространственной решетке со столь большим межатомным расстоянием, что их взаимодействием друг с другом можно пренебречь. Ясно, что энергетические уровни электронов в таком гипотетическом кристалле будут такие же, как и у изолированного атома. Разница будет только в том, что теперь каждому уровню энергии соответствует в N раз больше различных электронных состояний, чем в одном атоме.

Будем постепенно уменьшать межатомное расстояние. По мере сближения все более и более существенным становится взаимодействие между атомами, которое сказывается на уровнях энергии электронов. Равновесное расстояние между атомами в кристалле приблизительно таково, что электронные оболочки, соответствующие внешним (ва-

лентным) электронам, приходят в соприкосновение. При этом электроны внутренних электронных оболочек, размеры которых малы по сравнению с межатомным расстоянием в кристалле, почти не «чувствуют» поля, создаваемого соседними атомами, и их состояние в кристалле почти не отличается от состояния в изолированном атоме.

Наиболее сильному возмущающему воздействию соседних атомов подвергаются самые удаленные от ядра валентные электроны. В результате каждый уровень энергии валентных электронов расщепляется на большое число близко расположенных уровней, которые можно рассматривать как квазинепрерывную зону разрешенных значений энергии электронов.

Так как состояния исходной системы удаленных друг от друга атомов изменяются при их сближении непрерывным образом, то число различных состояний в кристалле должно быть таким же, как и в исходной системе. Поэтому число уровней в каждой зоне равно полному числу атомов в кристалле N. Однако ширина разрешенной энергетической зоны не зависит от полного числа атомов в кристалле, а определяется только межатомным расстоянием. При увеличении числа атомов в кристалле возрастает лишь густота энергетических уровней в пределах разрешенной зоны. Так как число атомов в кристалле велико ($N \sim 10^{23}$), то энергетический спектр электронов в пределах разрешенной зоны можно считать практически непрерывным.

Нас интересуют главным образом энергетические зоны, соответствующие внешним, т. е. валентным, электронам атомов, ибо именно они определяют большинство наблюдаемых макроскопических свойств кристаллов. Если уровень энергии валентного электрона атома в основном состоянии был заполнен электронами целиком, т. е. были заняты все различные состояния с данным значением энергии, то и в кристалле соответствующая энергетическая зона будет заполнена полностью. В этом случае зона называется валентной. Она отделена от расположенной выше разрешенной зоны, возникшей из наинизшего возбужденного уровня энергии атома, некоторым энергетическим интервалом, называемым запрещенной зоной. Когда кристалл находится в основном состоянии, все разрешенные зоны, расположенные выше валентной, пусты — в соответствующих этим зонам состояниях электронов нет. Диа-

грамма энергетических зон кристалла схематически показана на рис. 9.1. Заполненные электронами уровни энергии заштрихованы.

Такое заполнение энергетических зон электронами характерно для диэлектриков. В полностью заполненной валентной зоне диэлектриков свободных уровней энергии нет. Поэтому под действием приложенного к кристаллу электрического поля электрон в заполненной зоне не может изменить своего состояния. Этим объясняется отсутствие электропроводности у диэлектриков.

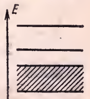


Рис. 9.1. Схема энергетических зон кристалла.

Если в кристалле запрещенная зона (т. е. расстояние от заполненной валентной зоны до ближайшей свободной зоны) невелика, так что в тепловом равновесии при конечной температуре часть электронов из валентной зоны в результате теплового возбуждения оказывается в свободной зоне, то кристалл представляет собой полупроводник. Находящиеся в почти пустой зоне электроны под действием внешнего электрического поля могут изменять свое состояние, т. е. ускоряться. Это означает, что в таком кристалле электрическое поле создает ток. Поэтому находящиеся в почти пустой зоне электроны называют электронами проводимости, а саму зону — зоной проводимости. Так как с ростом температуры число электронов проводимости увеличивается, электропроводность полупроводников растет с температурой.

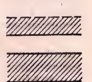


Рис. 9.2. В металле зона проводимости заполнена электронами частично.

Если уровень энергии валентного электрона в основном состоянии изолированного атома был заполнен частично, т. е. только часть разрешенных состояний с данной энергией занята электронами, то и при образовании кристалла из таких атомов соответствующая энергетическая зона будет заполнена лишь частично (рис. 9.2). В этом случае в зоне имеются не занятые электронами состояния, т. е. возможен электрический ток под действием приложенного поля. Так как число электронов в этой частично заполненной зоне проводимости очень велико — не меньше, чем число

атомов в кристалле,— то проводимость такого кристалла велика. Это — металлы.

Характерное для металлов заполнение зон может получиться и в том случае, когда у изолированного атома уровень энергии валентных электронов заполнен целиком, но при сближении атомов в кристалл происходит настолько сильное расщепление уровней, что верхняя целиком заполненная зона и соседняя с ней пустая зона начинают перекрываться. В этом случае наименьшее значение энергии кристалла соответствует такому заполнению энергетических зон, когда часть электронов с верхних уровней заполненной зоны переходит на нижние уровни пустой зоны (рис. 9.3).

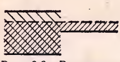


Рис. 9.3. Распределение электронов по энергетическим уровням в случае перекрытия свободной и заполненной зон.

Таким образом, деление твердых тел на диэлектрики, полупроводники и металлы, основанное в первую очередь на различии их электрического сопротивления, имеет под собой глубокую основу, связанную с различием их электронной структуры. Различный характер заполнения энергетических зон в диэлектриках и металлах приводит к исключительно большому различию их сопротивлений. Так, у чистых металлов при низких температурах удельное сопротивление может быть всего лишь 10^{-8} Ом·см, в то время как у диэлектриков с достаточно чистой поверхностью оно может достигать 10^{22} Ом·см. Таким образом, сопротивление диэлектриков и металлов может отличаться в 10^{30} раз!

Различие между проводниками и диэлектриками проявляется не только в электрическом сопротивлении, но и во многих других свойствах, например оптических: чистые однородные диэлектрические кристаллы прозрачны в видимой области спектра, а металлы — нет. Зато они хорошо отражают видимый свет.

Рассмотрим теперь подробнее стационарные состояния электронов, соответствующие какой-либо разрешенной энергетической зоне. Поскольку входящие в состав твердого тела атомы образуют правильную кристаллическую решетку, то эффективное поле, действующее на какой-либо электрон со стороны ядер и всех остальных электронов,

имеет пространственно периодический характер. Квантовомеханическое решение задачи о движении электрона в поле периодического потенциала приводит к следующим результатам. Стационарные состояния электрона в таком поле во многом напоминают состояния свободного электрона. Состояние свободной частицы характеризуется определенным значением импульса $\boldsymbol{p}$, поскольку для свободной частицы импульс является сохраняющейся величиной. Так как импульс имеет строго определенное значение, то вследствие соотношений неопределенностей Гейзенберга координаты электрона не имеют определенного значения: в таком состоянии электрон как бы размазан по всему пространству в том смысле, что вероятность обнаружить его в любом месте одинакова.

Как и у свободной частицы, состояния электрона в периодическом поле характеризуются вектором $\boldsymbol{p}$, который, в отличие от импульса свободной частицы, изменяется в некоторой ограниченной области, размер которой зависит от расстояния между атомами в кристалле. Как и свободная частица, электрон в стационарном состоянии в зоне не локализован, т. е. с равной вероятностью может быть обнаружен вблизи любого узла решетки. Оказывается, что зависимость энергии от импульса $E(\boldsymbol{p})$ для электрона в разрешенной зоне такая же, как и у свободной частицы, с той только разницей, что масса свободного электрона должна быть заменена на некоторую эффективную массу m^*:

$$E(\boldsymbol{p})=\frac{p^2}{2m^*}. \tag{9.1}$$

Энергия $E(\boldsymbol{p})$ отсчитывается от дна соответствующей зоны. Эффективная масса может быть как больше, так и меньше массы свободного электрона m. Более того, она может быть анизотропной, если свойства действующего на электрон периодического поля различны по разным направлениям.

В физике твердого тела понятие эффективной массы вводится для того, чтобы максимально приблизить описание движения электрона в разрешенной зоне к движению свободного электрона.

Как и свободная частица, электрон в кристалле под действием постоянного электрического поля движется равноускоренно. Иными словами, идеально регулярное

периодическое потенциальное поле не оказывает сопротивления электрическому току. Закон Ома в идеальном кристалле не может выполняться, так как он имеет место только тогда, когда движение зарядов в электрическом поле происходит с постоянной средней скоростью. Сопротивление электрическому току в реальном кристалле обусловлено не потенциальными барьерами периодического поля, а отступлениями поля реального кристалла от строгой периодичности либо за счет тепловых колебаний атомов, образующих кристалл, либо за счет различного рода дефектов решетки — примесей, вакансий и т. п.

§ 10. Электронные свойства простых металлов

Многие макроскопические свойства металлов, обусловленные наличием в них электронов проводимости, невозможно объяснить на основе классической статистической механики. Объяснение поведения коллектива электронов проводимости требует обязательного привлечения квантовых закономерностей.

Рассмотрим подробнее свойства электронов проводимости в металлах, считая для простоты, что их эффективная масса изотропна, а число электронов проводимости в образце равно числу атомов металла N. Именно так обстоит дело, например, у щелочных металлов. Хотя поведение отдельного электрона проводимости в кристалле очень сходно с поведением классической свободной частицы, свойства коллектива таких электронов резко отличаются от свойств классического газа. Причина этому — принцип Паули, согласно которому в каждом квантовом состоянии может находиться не более одного электрона. Принцип Паули приводит к тому, что даже при абсолютном нуле температуры электроны не могут находиться в состоянии покоя, так как нельзя все электроны поместить в состояние с равным нулю импульсом. Оказывается, что уровню энергии с нулевым импульсом соответствуют только два различных квантовых состояния. Поэтому только два из N электронов могут иметь равные нулю импульс и энергию. Всем остальным электронам придется занимать состояния с отличными от нуля импульсом и энергией. При температуре $T=0$ К система электронов будет иметь наименьшую возможную энергию, если все состояния с импульсами,

меньшими некоторого предельного значения, называемого импульсом Ферми p_F, будут заняты, а все состояния с бо́льшими импульсами — пусты. Функция распределения электронов по состояниям $f(\boldsymbol{p})$, дающая вероятность заполнения данного состояния с импульсом $\boldsymbol{p}$, будет, следовательно, при $T=0$ К равна единице при $|\boldsymbol{p}|\leqslant p_F$ и равна нулю при $|\boldsymbol{p}|>p_F$ (рис. 10.1). Величина импульса Ферми зависит от объемной концентрации электронов в зоне проводимости $n=N/V$, где V — объем образца. Характер этой зависимости можно установить на основании соотношений неопределенностей Гейзенберга. Пусть образец металла имеет форму прямоугольного параллелепипеда со сторонами L_x, L_y, L_z. Для электрона, находящегося внутри образца, наибольшая неопределенность в значении x-координаты равна размеру образца в этом направлении; поэтому наименьшая неопределенность его x-компоненты импульса Δp_x определяется соотношением

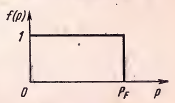


Рис. 10.1. Функция распределения электронов при абсолютном нуле температуры.

$$\Delta p_x \sim \frac{h}{L_x}. \qquad (10.1)$$

Такие же соотношения справедливы и для Δp_y, Δp_z. Поэтому для произведения неопределенностей имеем

$$\Delta p_x \Delta p_y \Delta p_z \sim \frac{h^3}{L_x L_y L_z} = \frac{h^3}{V}. \qquad (10.2)$$

Поскольку Δp_x есть наименьшая неопределенность значения p_x, совместимая с условием пребывания электрона внутри образца, то Δp_x есть та наименьшая величина, на которую следует изменить значение p_x для того, чтобы можно было считать, что электрон находится уже в другом квантовом состоянии. Поэтому произведение $\Delta p_x \Delta p_y \Delta p_z$ дает объем ячейки в пространстве импульсов, при изменении импульса электрона в пределах которой его состояние не меняется. Это означает, что N электронов в соответствии с принципом Паули занимают в пространстве импульсов объем, пропорциональный $N\Delta p_x \Delta p_y \Delta p_z$. В соответствии

с (10.2)

$$N\,\Delta p_x\,\Delta p_y\,\Delta p_z \sim N\frac{h^3}{V} = h^3 n. \qquad (10.3)$$

С другой стороны, этот объем равен объему шара радиуса p_F с центром в начале координат: $^4/_3\pi p_F^3$. Поэтому $p_F^3 \sim h^3 n$, откуда

$$p_F \sim h n^{1/3}. \qquad (10.4)$$

Наибольшая энергия, которую может иметь электрон проводимости в металле при $T=0$ K,— это энергия электрона с $p=p_F$:

$$E_F = \frac{p_F^2}{2m^*} \sim \frac{h^2 n^{2/3}}{m^*}. \qquad (10.5)$$

Интересно сравнить энергию Ферми E_F с характерной тепловой энергией kT. Для типичных металлов концентрация электронов проводимости составляет $\sim 10^{21}$—10^{23} см$^{-3}$, а эффективная масса m^* близка к массе свободного электрона. Поэтому по формуле (10.5) легко подсчитать, что E_F оказывается порядка нескольких электронвольт. Легко убедиться, что тепловая энергия kT становится сравнимой с энергией Ферми только при температурах порядка 50 000 K! Для всех металлов при всех температурах вплоть до точки плавления $kT \ll E_F$.

Нетрудно подсчитать скорость электрона, обладающего энергией E_F. Она оказывается равной примерно 10^8 см/с. Таким образом, вследствие принципа Паули электроны проводимости в металле при абсолютном нуле температуры отнюдь не покоятся.

При повышении температуры от абсолютного нуля до некоторого значения T средняя энергия каждой частицы, согласно представлениям классической статистической механики, должна возрасти на величину порядка kT. Но для подчиняющихся принципу Паули электронов в металле это не так. При температуре абсолютного нуля электроны проводимости уже имеют энергию порядка E_F, которая всегда много больше kT. Эта энергия не связана с тепловым движением электронов. При повышении температуры от 0 K до T отнюдь не каждый электрон может увеличить свою энергию на величину $\sim kT$, так как для большинства электронов все состояния, отстоящие от данного по энергии

на величину порядка kT, заняты другими электронами. Увеличить свою энергию смогут только электроны, занимающие состояния в интервале $\sim kT$ около E_F. Поэтому тепловое движение при конечной температуре приводит к размытию ступеньки функции распределения f в узком интервале значений энергии $\sim kT$ (рис. 10.2). Основная часть электронов будет занимать те же состояния, что и при абсолютном нуле. О таких свойствах функции распределения говорят как о сильном вырождении электронного газа в металлах.

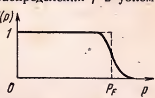


Рис. 10.2. Температурное размытие края функции распределения электронов в металле.

Отличие функции распределения электронов проводимости в металле от максвелловской функции распределения классического газа объясняет особенности наблюдаемых электронных свойств металлов. Сравним, например, теплоемкости классического одноатомного газа и вырожденного газа электронов проводимости. В состоянии теплового равновесия в одноатомном классическом газе средняя энергия хаотического движения одной частицы равна $^3/_2kT$. Поэтому внутренняя энергия всего газа из N частиц равна $^3/_2NkT$, а его теплоемкость при постоянном объеме равна $^3/_2Nk$ и не зависит от температуры.

В вырожденном электронном газе в тепловом движении может участвовать лишь часть электронов порядка kT/E_F, поскольку именно такая доля полного их числа N содержится в интервале шириной порядка kT вблизи энергии Ферми E_F. Каждый из этих электронов приобретает дополнительную энергию порядка kT, и, следовательно, полная энергия хаотического теплового движения электронов U по порядку величины равна

$$U \approx N \frac{kT}{E_F} kT. \qquad (10.6)$$

Отсюда для теплоемкости электронного газа получаем

$$C_V \approx Nk \frac{kT}{E_F}. \qquad (10.7)$$

Поскольку $kT/E_F \ll 1$ (например, при комнатной температуре это отношение равно всего лишь 1/200), то теплоем-

кость вырожденного газа (10.7) примерно на два порядка меньше теплоемкости классического газа. Поэтому при комнатной температуре молярная теплоемкость металлов практически целиком определяется вкладом решетки и почти не отличается от молярной теплоемкости диэлектриков, в которых электроны проводимости отсутствуют.

Другое важное отличие теплоемкости электронного газа (10.7) заключается в том, что она пропорциональна температуре, в то время как теплоемкость классического газа от температуры не зависит.

Наряду со свойствами коллектива электронов проводимости, которые он проявляет в состоянии равновесия, большой интерес представляет изучение поведения электронов проводимости в металле в неравновесном состоянии, когда они движутся под действием приложенных внешних полей. Такие процессы носят название явлений переноса. Примерами таких явлений могут служить теплопроводность, электропроводность, эффект Холла, термоэлектрические явления и т. д.

Остановимся подробнее на одном из наиболее важных неравновесных процессов — прохождении электрического тока в металле, описываемом законом Ома. Закон Ома является, пожалуй, одним из самых ранних экспериментальных открытий в области электрического тока в веществе. Согласно этому закону плотность электрического тока пропорциональна напряженности приложенного электрического поля: $\boldsymbol{j}=\sigma\boldsymbol{E}$. Это уравнение описывает наблюдаемое поведение очень многих веществ в широком диапазоне условий.

Закон Ома нельзя вывести только из законов электродинамики, описывающих электрическое поле. Его можно получить на основе изучения процессов, происходящих в веществе при приложении электрического поля. Интересно отметить, что в разных веществах эти процессы могут сильно различаться, но, независимо от характера процессов рассеяния носителей, связь между плотностью тока и напряженностью поля при не слишком сильных полях всегда оказывается линейной.

В металлах сопротивление электрическому току обусловлено рассеянием электронов проводимости на тепловых колебаниях решетки и хаотически расположенных дефектах кристалла — в идеальной решетке сопротивление от-

сутствует. Феноменологически описать процессы рассеяния электронов проводимости можно путем введения так называемого времени релаксации. Будем рассуждать следующим образом. В состоянии равновесия функция распределения электронов по импульсам сферически симметрична. Поэтому средняя скорость электронов равна нулю и никакого направленного переноса заряда, т. е. электрического тока, нет. Приложенное электрическое поле вызывает искажение вида функции распределения. В отличие от изотропного теплового воздействия, которое могло только размыть край распределения, не нарушая его сферической симметрии, направленное действие внешнего поля вызывает сдвиг всего распределения. Для такого сдвинутого распределения средняя скорость уже отлична от нуля, что и означает наличие тока в образце.

Пусть в некоторый момент времени внешнее электрическое поле выключается. Очевидно, что система электронов будет возвращаться в состояние теплового равновесия. Предположим, что скорость приближения к равновесию в каждый момент пропорциональна величине отклонения от равновесия. Тогда и изменение средней скорости направленного движения электронов будет происходить по такому же закону. Так как в равновесии средняя скорость равна нулю, то изменение скорости происходит пропорционально самой средней скорости:

$$\frac{d\langle \boldsymbol{v}\rangle}{dt} = -\frac{\langle \boldsymbol{v}\rangle}{\tau}. \qquad (10.8)$$

Поскольку решение этого уравнения имеет вид

$$\langle \boldsymbol{v}(t)\rangle = \boldsymbol{v}_0 e^{-t/\tau}, \qquad (10.9)$$

то τ есть то время, в течение которого скорость направленного движения убывает в e раз. Оно называется временем релаксации.

Закон изменения средней скорости $\langle \boldsymbol{v}(t)\rangle$ направленного движения электронов проводимости (10.9) имеет в точности такой же вид, что и уравнение движения тела в вязкой среде под действием силы трения, пропорциональной скорости: $\boldsymbol{F}_{\text{тр}} = -\frac{m}{\tau}\boldsymbol{v}$. Это наводит на мысль, что взаимодействие электронов проводимости с колебаниями решетки и примесями, приводящее к описываемой уравнением (10.9)

релаксации средней скорости, можно учесть путем введения некоторой эффективной, непрерывно действующей на электрон силы трения $F_{тр} = -\frac{m}{\tau} v$. При наличии постоянного электрического поля E на электрон, кроме силы трения, действует еще сила $-eE$, где $-e$ — заряд электрона. В стационарном состоянии электрон движется с такой постоянной скоростью, при которой эти силы уравновешиваются: $-eE = \frac{m}{\tau} v$. Такой же будет и средняя скорость $\langle v \rangle$ движения электронов в металле под действием поля E: $\langle v \rangle = -\frac{e\tau}{m} E$. Плотность тока j равна $-en \langle v \rangle$. Подставляя сюда значение $\langle v \rangle$, находим

$$j = \frac{ne^2\tau}{m} E = \sigma E.$$

Это и есть закон Ома. Вычисление времени релаксации τ и, следовательно, удельной электропроводности σ требует детального рассмотрения механизмов рассеяния электронов проводимости.

§ 11. Плазма и электроны в металлах

При изучении электронных свойств твердых тел, таких, как теплоемкость и электропроводность, систему свободных электронов можно было рассматривать как совокупность невзаимодействующих частиц. Однако существует целый ряд явлений, в которых взаимодействие электронов играет определяющую роль. В таких случаях говорят о проявлении электронами плазменных свойств.

Вообще, плазмой называют систему, состоящую из большого числа подвижных частиц, по крайней мере часть которых обладает электрическим зарядом. Термин «плазма» широко используется в современной физике. Этот термин применяют к ионизированному газу при таких условиях, когда силами взаимодействия составляющих его частиц нельзя пренебрегать. Плазмой называют и электронный газ в металлах и полупроводниках.

Плазма — это наиболее распространенное состояние вещества в природе. Звезды представляют собой гигантские сгустки горячей плазмы. Внешний ионизированный слой земной атмосферы, радиационные пояса, некоторые типы

кометных хвостов, наконец, пламя мартеновской печи — все это примеры систем, являющихся плазмой.

Отличительные свойства плазмы связаны с тем, что в ее состав входят частицы, обладающие электрическим зарядом. Новые, необычные для нейтрального газа свойства плазмы обусловлены тем чрезвычайно сильным воздействием, которое оказывают электрические и магнитные поля на движение заряженных частиц. На нейтральные частицы электрические и магнитные поля оказывают гораздо меньшее воздействие. В газе нейтральных частиц информация о локальном изменении состояния, например об увеличении концентрации частиц в каком-либо месте, передается лишь в результате столкновений частиц. В плазме картина иная. В отличие от нейтральных частиц, которые взаимодействуют друг с другом только на малых расстояниях, заряженные частицы взаимодействуют с помощью дальнодействующих кулоновских сил. При локальном изменении состояния в плазме возникают электрическое и магнитное поля, которые действуют на всю плазму в целом. В результате в плазме возникают коллективные движения частиц — колебания и волны. Скорость передачи информации о локальных возмущениях определяется скоростью распространения электромагнитных волн в плазме. Именно наличие таких специфических коллективных процессов в плазме — волн, в которых происходят колебания как частиц плазмы, так и сопровождающих движение заряженных частиц электромагнитных полей,— и позволяет говорить о плазме как о четвертом состоянии вещества.

Газоразрядная плазма, как правило, представляет собой смесь трех компонент: свободных электронов, положительных ионов и нейтральных атомов или молекул. Электроны — это наиболее подвижная часть плазмы, и именно с движением электронов связаны ее наиболее интересные свойства. Ионы же вследствие гораздо большей массы ведут себя более «пассивно», благодаря чему во многих случаях можно вообще пренебрегать их движением и рассматривать ионную часть плазмы как неподвижный положительный фон, на котором происходит движение электронов. Такое приближение однокомпонентной плазмы тем более оправдано для электронов в металлах, где движение ионов вообще ограничено колебаниями вблизи положений равновесия.

Одним из важнейших свойств плазмы является ее стремление к сохранению равенства плотностей электрического заряда положительных и отрицательных частиц. В самом деле, при высокой плотности заряженных частиц в плазме даже малое пространственное разделение положительных и отрицательных зарядов привело бы к появлению очень сильных электрических полей, стремящихся восстановить локальное нарушение электронейтральности. Поэтому в среднем (в достаточно большом объеме или за достаточно большой промежуток времени) плазма должна быть почти нейтральной, или, как говорят, квазинейтральной.

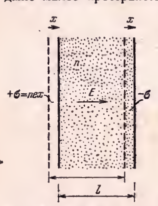


Рис. 11.1. К выводу формулы для частоты плазменных колебаний.

Оценим величину объема и промежутка времени, в которых выполняется квазинейтральность плазмы. Представим себе, что в плоском слое однородной нейтральной в целом плазмы все электроны сместились на расстояние x в одном и том же направлении (рис. 11.1). Возникающее в результате такого смещения электронов результирующее распределение зарядов будет таким же, как и в плоском конденсаторе. Электрическое поле в плазме определяется плотностью заряда на «обкладках» такого «конденсатора». При смещении электронов в слое толщиной l (рис. 11.1) нарушение нейтральности происходит только в тонких областях толщиной x вблизи границ слоя: слева образуется избыток положительного заряда, справа — отрицательного. Если концентрацию электронов в нейтральной плазме обозначить через n, то при смещении всех электронов в слое на расстояние x заряд σ на единице площади «обкладки» будет равен enx. Поэтому напряженность поля E в «конденсаторе» будет равна

$$E=4\pi\sigma=4\pi enx. \qquad (11.1)$$

Действующая на каждый электрон в слое сила $\boldsymbol{F}=-e\boldsymbol{E}$ будет пропорциональна смещению электрона x и, как видно из рис. 11.1, направлена в сторону, противоположную

смещению:

$$F=-4\pi ne^2x. \qquad (11.2)$$

Поэтому электроны будут совершать гармонические колебания, частота которых ω_p определяется выражением

$$\omega_p^2=\frac{4\pi ne^2}{m}, \qquad (11.3)$$

где m — масса электрона.

Эта частота ω_p — одна из важнейших характернстнк плазмы. Ее называют плазменной частотой, а сами колебания — плазменными или ленгмюровскимн, по именн амернканского физика Ленгмюра, впервые исследовавшего этн колебания.

Таким образом, в результате разделения зарядов в плазме возникают электрические поля, вызывающие колебания частиц. Этн колебания стремятся восстановить квазинейтральность плазмы. Ясно, что заметить отклонения плазмы от квазинейтральности можно только на протяженин времени, малого по сравнению с периодом плазменных колебаний. В среднем (за много периодов колебаний) плазма ведет себя как квазннейтральная среда. Период плазменных колебаний $T_p=2\pi/\omega_p$ — это характерный временной масштаб разделения зарядов в плазме.

Теперь оценим пространственный масштаб разделения зарядов в плазме. Очевидно, что отклонение плазмы от квазинейтральности может проявляться только в области, размер которой меньше амплнтуды плазменных колебаний. Но эта амплнтуда не может быть сколь угодно большой из-за теплового движения электронов, которое приводит к разрушению электрнческих полей, связанных с плазменными колебаниями. Возникающее при плазменных колебаннях пространственное разделение зарядов «смазывается» в результате теплового движения электронов. Максимальное расстояние, на котором еще возможно разделение зарядов, не может превышать расстояния, проходимого электроном с характерной тепловой скоростью за время существования пространственного разделения зарядов. Поскольку это время порядка периода плазменных колебаний, то для пространственного масштаба разделения зарядов в плазме получаем следующее выражение:

$$r_0\sim\sqrt{\frac{kT}{m}}\frac{1}{\omega_p}=\sqrt{\frac{kT}{4\pi ne^2}}. \qquad (11.4)$$

Эта величина носит название дебаевского радиуса экранирования, по имени английского физика Дебая, который впервые ввел ее при изучении экранировки кулоновского взаимодействия заряженных частиц. На расстоянии дебаевского радиуса происходит экранирование кулоновского поля любого заряда. Причиной этого экранирования является преимущественная группировка вокруг любого заряда заряженных частиц противоположного знака. Кулоновские силы стремятся максимально приблизить к внесенному в плазму пробному заряду заряженные частицы противоположного знака, а хаотическое тепловое движение препятствует этому. В результате вокруг пробного заряда возникает пространственно неоднородное распределение электронов, приводящее к нейтрализации всей системы на расстоянии порядка дебаевского радиуса r_0.

Полностью ионизированная однокомпонентная квазинейтральная плазма в состоянии термодинамического равновесия характеризуется четырьмя параметрами: зарядом частиц e, их массой m, концентрацией n и температурой T. Нетрудно убедиться, что из этих величин можно составить только один независимый безразмерный параметр γ:

$$\gamma = \frac{e^2 n^{1/3}}{kT}. \qquad (11.5)$$

Постоянная Больцмана k в этом выражении появляется потому, что в системе единиц СГС температуру, характеризующую среднюю энергию теплового движения, необходимо выразить в эргах. Другими словами, если через T обозначена температура в градусах Кельвина, то в эргах ей соответствует величина kT.

Безразмерный параметр γ имеет простой физический смысл: с точностью до множителя порядка единицы он равен отношению потенциальной энергии электростатического взаимодействия двух частиц плазмы, находящихся на расстоянии $n^{-1/3}$ друг от друга, равном среднему расстоянию между частицами, к средней энергии теплового движения частиц. Если этот параметр $\gamma \ll 1$, то потенциальная энергия мала по сравнению с кинетической и плазма по своим термодинамическим свойствам близка к идеальному газу. Таким образом, классическая плазма, как видно из формулы (11.5), является тем более идеальной, чем выше ее температура и чем ниже концентрация частиц в ней.

До сих пор молчаливо предполагалось, что к частицам, из которых состоит плазма, можно применять законы классической физики. В газоразрядной и тем более в ионосферной плазме это действительно так, ибо расстояние между электронами много больше их дебройлевской длины волны. А как обстоит дело для плазмы, представляющей собой электроны в металлах или полупроводниках? В этом случае поведение электронов описывается квантовой механикой, и мы должны ожидать, что в формулах, выражающих свойства такой плазмы, наряду с параметрами e, m, n и kT появится постоянная Планка h. Теперь наряду с безразмерным параметром γ, даваемым формулой (11.5), появляется еще один независимый безразмерный параметр Γ:

$$\Gamma = \frac{h^2 n^{2/3}}{mkT}. \tag{11.6}$$

Физический смысл этого параметра легко уяснить, обращаясь к формуле (10.5): Γ представляет собой отношение энергии Ферми E_F к характерной тепловой энергии kT. Если $\Gamma \ll 1$, т. е. $E_F \ll kT$, то электронный газ по своим свойствам близок к рассмотренной выше классической плазме. Так бывает в полупроводниках при не слишком больших концентрациях электронов проводимости и достаточно высоких температурах.

Как мы видели в предыдущем параграфе, для электронов в металлах при всех температурах вплоть до точки плавления, $E_F \gg kT$ и, следовательно, $\Gamma \gg 1$. Для такой плазмы потенциальную энергию кулоновского взаимодействия электронов $e^2 n^{1/3}$ следует сравнивать не с кинетической энергией теплового движения kT, а с энергией Ферми $E_F = p_F^2/2m \sim \sim h^2 n^{2/3}/m$. Степень близости электронной плазмы в металлах к идеальному газу характеризуется безразмерным параметром, который принято обозначать r_s:

$$r_s = \frac{e^2 n^{1/3}}{h^2 n^{2/3}/m} = \frac{e^2 m}{h^2 n^{1/3}}. \tag{11.7}$$

Нетрудно видеть, что параметр r_s пропорционален отношению среднего расстояния между электронами $n^{-1/3}$ к боровскому радиусу $a_0 = \hbar^2/(me^2)$. Электронный газ в металле был бы близок к идеальному газу при $r_s \ll 1$. Из формулы (11.7) видно, что, в противоположность класси-

ческой плазме, электронный газ тем ближе к идеальному газу, чем выше его концентрация. Конечно, при увеличении концентрации электронов n потенциальная энергия их взаимодействия возрастает. В случае классической плазмы, где кинетическая энергия не зависит от концентрации, степень близости плазмы к идеальному газу при этом убывала. Для электронов в металле, как мы видели, вследствие принципа Паули кинетическая энергия растет с увеличением концентрации, как $n^{2/3}$. Поэтому, несмотря на увеличение потенциальной энергии, которая растет, как $n^{1/3}$, отношение потенциальной энергии к кинетической убывает. Другими словами, относительная роль взаимодействия при увеличении концентрации электронов становится все меньше и меньше. Оказывается, что для реальных металлов параметр r_s больше единицы, т. е. концентрация электронов все-таки недостаточно высока для того, чтобы электронный газ можно было считать идеальным. Для электронов в металле более подходящим является название «электронная жидкость», чем «электронный газ». Тем не менее последовательная теория электронной жидкости металлов показала, что многие свойства, полученные в модели свободных электронов, качественно остаются справедливыми и при учете электронного взаимодействия.

Частота плазменных колебаний ω_p представляет собой граничную частоту для электромагнитных волн, которые могут распространяться в плазме. Волны с частотой, меньшей ω_p, не могут проникать в плазму, так как низкочастотные электромагнитные поля таких волн экранируются заряженными частицами плазмы. Падающая на границу плазмы волна при $\omega < \omega_p$ отражается от границы. Если же частота электромагнитной волны выше плазменной частоты, то такая волна проникает в плазму. Концентрация свободных электронов в щелочных металлах такова, что частота плазменных колебаний соответствует ультрафиолетовой области спектра. Поэтому в ультрафиолетовой области щелочные металлы прозрачны, хотя в видимой и инфракрасной областях спектра хорошо отражают падающее на них излучение. Для ионосферной плазмы граница прозрачности попадает в диапазон метровых радиоволн.

Приложение. СИСТЕМЫ ЕДИНИЦ

Любое измерение заключается в сравнении измеряемой величины с другой, однородной с ней величиной, принятой за единицу.

В принципе единицы для всех величин можно выбрать совершенно независимо друг от друга. Однако это практически неудобно, так как тогда во всех уравнениях физических законов, выражающих связь между различными величинами, появятся численные коэффициенты. Кроме того, пришлось бы для каждой физической величины вводить свой эталон. Поэтому основной особенностью современных единиц является то, что между единицами различных величин имеются определенные соотношения. Эти соотношения устанавливаются теми физическими законами или определениями, которыми связаны между собой измеряемые величины. Например, единица скорости выбирается так, что она выражается через единицы длины и времени. При таком выборе единицы скорости используется определение скорости.

Это означает, что при построении определенной системы единиц для нескольких произвольно выбираемых физических величин единицы устанавливаются независимо друг от друга и называются основными. Единицы для остальных величин выражаются через основные и носят название производных единиц. Число основных единиц и сам их выбор в разных системах единиц могут быть различными. Например, для электрических и магнитных измерений в системе СГС в качестве основных выбраны три единицы: единицы длины (L), времени (T) и массы (M), а в международной системе единиц СИ в качестве основных выбираются четыре: единицы длины, времени, массы и силы тока (I).

Кроме произвола в выборе физических величин, единицы которых принимаются за основные, и произвола в выборе масштаба (размера) этих единиц, имеется еще произвол в выборе коэффициентов пропорциональности в формулах, которыми вводятся производные единицы. Проиллюстрируем это на примере единицы площади. Выбрав в качестве единицы длины метр, можно в качестве единицы площади выбрать либо квадратный метр — площадь квадрата, сторона которого

равна метру,— либо «круглый» метр — площадь круга, диаметр которого равен метру. В первом случае площадь квадрата со стороной l выражается формулой $S=l^2$, а площадь круга с диаметром l — формулой $S=\pi l^2/4$. Во втором случае более простая формула получается для площади круга: $S=l^2$, в то время как формула для площади квадрата будет содержать π: $S=4l^2/\pi$.

Рассмотренные возможности введения производных единиц площади, отличающихся численным коэффициентом, основывались на одной и той же геометрической закономерности, связывающей площади подобных фигур с их линейными размерами. Но при введении производной единицы для какой-либо величины, кроме упомянутого произвола в выборе численного коэффициента, имеется еще произвол в выборе физического закона, с помощью которого устанавливается связь производных единиц с основными. Например, единица силы обычно устанавливается с помощью второго закона Ньютона $F=ma$. В этом случае выражение единицы силы через основные единицы, т. е. ее размерность, имеет вид

$$[F]=MLT^{-2}. \tag{1}$$

Однако при тех же основных единицах (L, M, T) для установления производной единицы силы можно вместо второго закона Ньютона использовать закон всемирного тяготения, полагая в нем коэффициент пропорциональности безразмерным и равным, например, единице: $F=m_1m_2/r^2$. В этом случае за единицу силы принимается сила, с которой притягиваются друг к другу единичные точечные массы, находящиеся на единичном расстоянии друг от друга. Размерность силы при этом имеет вид

$$[F]=M^2L^{-2}. \tag{2}$$

При таком выборе единицы силы во втором законе Ньютона, разумеется, появится размерный коэффициент, подобно тому как при обычном выборе (на основе второго закона Ньютона) он появляется в законе всемирного тяготения и носит название гравитационной постоянной. Разобранный пример показывает, что размерность физической величины зависит от способа построения системы единиц.

Таким образом, мы видим, что при выборе способа построения системы единиц существует большой произвол. Однако на практике приходится считаться с целым рядом требований, которые существенно ограничивают этот произвол. Слишком большое число основных единиц было бы неудобно из-за появления размерных коэффициентов во многих физических формулах и необходимости установления большого числа эталонов. Слишком малое число основных единиц приводит к тому, что

построенные на них производные единицы оказываются неудобными для использования на практике. Практически используются системы, в которых число основных единиц колеблется от трех до семи.

При установлении основных единиц весьма важной является возможность создания таких эталонов, которые обеспечивали бы постоянство единицы и возможность ее воспроизведения, а также восстановления эталона в случае его утраты. Самый надежный способ решения этой задачи — поручить «хранение» эталонов самой природе. Поэтому принятый в настоящее время эталон длины связывает единицу длины с длиной волны излучения узкой оранжевой спектральной линии изотопа криптона-86. По определению один метр содержит точно 1 650 763,73 длин волн в вакууме этой спектральной линии. Аналогично, эталон времени основывается на периоде колебаний, происходящих в атоме изотопа цезия-133. По определению единица времени — секунда — содержит 9 192 631 770 периодов этих колебаний. Атомы одного и того же изотопа тождественны, поэтому при указанном выборе эталонов длины и времени природа предоставляет в наше распоряжение практически неограниченное число совершенно идентичных «линеек» и «часов». Для эталона массы пока не удается использовать массу какой-либо атомной частицы, так как точность определения атомных масс уступает точности измерения массы при взвешивании. Эталоном массы служит платино-иридиевая гиря, хранящаяся в Международном бюро мер и весов в Севре, под Парижем.

Перейдем к подробному описанию двух наиболее употребительных систем единиц — СИ и так называемой системы Гаусса, которая получила широкое распространение в научной литературе по физике. Единицы механических величин в этих двух системах отличаются только масштабом, так как основные единицы в них выбраны на основе одних и тех же физических величин — длины, времени и массы. Поэтому вид всех формул и уравнений, выражающих физические законы и определения, в механике одинаков в обеих системах.

Иначе обстоит дело в электродинамике. Для электрических величин единицы гауссовой системы совпадают с единицами системы СГСЭ. Единица заряда является производной и выражается через основные с помощью закона Кулона, коэффициент пропорциональности в котором выбирается безразмерным и равным единице: $F=q_1q_2/r^2$. Размерность заряда получается равной

$$[q]=M^{1/2}L^{3/2}T^{-1}. \qquad (3)$$

Единица заряда в системе СГСЭ не имеет специального названия. Все остальные электрические величины имеют единицы, выражающиеся через эту единицу заряда, а тем самым и через основные единицы.

Например, размерность силы тока есть

$$[I]=M^{1/2}L^{3/2}T^{-2}. \qquad (4)$$

Аналогичным образом вводятся производные единицы напряженности электрического поля, потенциала, емкости и т. д.

Единицы магнитных величин вводятся в гауссовой системе следующим образом. Рассмотрим магнитное поле, создаваемое прямолинейным бесконечным проводом, по которому течет ток I. Согласно закону Био—Савара—Лапласа элемент этого провода Δl создает в точке наблюдения, находящейся на расстоянии r от элемента, индукцию ΔB, равную

$$\Delta B=k\frac{I\Delta l\sin\alpha}{r^2},$$

где коэффициент k зависит от выбора единиц. Суммирование полей, создаваемых всеми элементами провода, дает для результирующей индукции поля в точке наблюдения выражение

$$B=k\frac{2I}{r}. \qquad (5)$$

Обнаружить магнитное поле можно по его действию на другой проводник с током. Если этот проводник расположить параллельно проводу, создающему магнитное поле, то действующая на него сила, согласно закону Ампера, будет пропорциональна индукции поля B, току в нем I' и его длине l:

$$F=k'I'Bl. \qquad (6)$$

Коэффициент k в формуле (5) может быть выбран произвольно, так как единица индукции поля B еще не установлена. Но после того, как этот коэффициент k в (5) выбран (тем самым выбрана и единица индукции B), коэффициент k' в формуле (6) уже не может выбираться произвольно, а должен определяться из эксперимента. Разумеется, можно поступить и наоборот: использовать уравнение (6) для введения единицы индукции поля, полагая $k'=1$, тогда коэффициент k в (5) будет определяться на опыте. В системе Гаусса поступают следующим образом. Выбирают коэффициент k в формуле (5) так, чтобы он равнялся коэффициенту k' в формуле (6).

Если подставить в формулу (6) индукцию B из (5), то для силы взаимодействия двух параллельных проводников с токами I и I', находящихся на расстоянии r друг от друга, получим следующее выражение:

$$F=kk'\frac{2II'l}{r}. \qquad (7)$$

В гауссовой системе $kk'=k^2$. Поскольку для всех величин, входящих в эту формулу, единицы уже выбраны, коэффициент k^2, как легко убедиться, имеет размерность $L^{-2}T^2$, т. е. обратную размерности квадрата скорости. Величина этого коэффициента должна определяться экспериментально по измерению силы взаимодействия двух параллельных проводников, находящихся на известном расстоянии, когда по ним протекают известные токи. Опыт показал, что численное значение k^2 равно $1/c^2$, где c — скорость света в вакууме:

$$c \approx 2{,}9979246 \times 10^{10} \text{ см/с}.$$

В гауссовой системе единиц закон Био—Савара—Лапласа и закон Ампера записываются в виде

$$\Delta B = \frac{1}{c}\,\frac{I\Delta l \sin\alpha}{r^2}, \quad F = \frac{1}{c}\,I'Bl.$$

На основании последней формулы устанавливается единица индукции магнитного поля — гаусс. Один гаусс — это индукция такого поля, которое действует на 1 см проводника с током в одну СГСЭ-единицу с силой, равной $1/c$ дин, если проводник расположен перпендикулярно линиям индукции магнитного поля.

В отличие от гауссовой системы единиц, где единица для силы тока является производной и выражается через основные единицы (L, M, T) с помощью закона Кулона, в системе СИ единица силы тока является основной. Эта единица выбрана следующим образом. В основу кладется закон взаимодействия параллельных токов, выражаемый формулой (7). За единицу силы тока, называемую ампером, принимается такой ток, при протекании которого по параллельным проводам, расположенным в вакууме на расстоянии 1 м друг от друга, сила взаимодействия, приходящаяся на 1 м длины провода, равна $2 \cdot 10^{-7}$ ньютонов. Наряду с метром, секундой и килограммом, ампер является четвертой основной единицей системы СИ. После того, как единица силы тока выбрана, можно установить размерность коэффициента kk' в формуле (7). Она оказывается равной ньютон/(ампер)^2, т. е. $MLT^{-2}I^{-2}$. Этот коэффициент kk' в системе СИ обозначается $\mu_0/4\pi$, а величина μ_0 называется магнитной постоянной. С помощью формулы (7) и определения ампера можно вычислить ее величину. Подставляя в формулу (7) $F=2\cdot10^{-7}$ Н, $I=I'=1$ А, $l=r=1$ м, $kk'=\mu_0/4\pi$, находим

$$\mu_0 = 4\pi\cdot10^{-7} \text{ Н/А}^2. \quad (8)$$

Численный множитель 4π вводится в формулу (8) для так называемой «рационализации» системы единиц. Благодаря этому коэффициент 4π исчезает из многих часто употребляемых формул (хотя при этом он

появляется в некоторых других формулах). Введение здесь множителя 4π совершенно аналогично рассмотренному выше примеру перехода от квадратных метров к «круглым» при измерении площади.

Подчеркнем, что численное значение магнитной постоянной μ_0 (8) получается как прямое следствие определения ампера, а не устанавливается на опыте, в отличие от коэффициента $1/c^2$ в гауссовой системе. Так получается потому, что в системе СИ единица силы тока является основной (выбранной произвольно), в то время как в гауссовой системе эта единица является производной.

Как мы видим, введение единицы силы тока — ампера — однозначно определяет только произведение коэффициентов k и k', входящих в формулы (5) и (6): $kk'=\mu_0/4\pi$. При этом еще остается произвол в выборе самих сомножителей.

В системе СИ полагают $k'=1$; тем самым для k получается значение, равное $\mu_0/4\pi$. В результате закон Ампера, описывающий действие магнитного поля на проводник с током I', в системе СИ записывается в виде

$$F=I'Bl. \tag{9}$$

Этот закон используется для установления единицы измерения индукции магнитного поля B. Эта единица носит название тесла и имеет размерность $MT^{-2}I^{-1}$:

$$[B]=\frac{\text{Н}}{\text{А}\cdot\text{м}}=\frac{\text{кг}}{\text{А}\cdot\text{с}^2}. \tag{10}$$

Таким образом, 1 тесла — это индукция такого поля, которое на 1 *м* проводника с током в 1 А действует с силой 1 Н, если проводник расположен перпендикулярно линиям индукции магнитного поля. Закон Био—Савара—Лапласа, на основе которого рассчитывается индукция магнитного поля, создаваемого проводниками с током, в системе СИ содержит размерный коэффициент $k=\mu_0/4\pi$:

$$\Delta B=\frac{\mu_0}{4\pi}\frac{I\Delta l\sin\alpha}{r^2}.$$

Закон Кулона, описывающий взаимодействие точечных зарядов, в системе СИ содержит размерный коэффициент, ибо единица заряда в СИ — 1 кулон — устанавливается независимо от закона Кулона на основании единицы силы тока — ампера: 1 Кл = 1 А·с. Размерный коэффициент в законе Кулона в системе СИ записывается в виде $1/(4\pi\varepsilon_0)$, где ε_0 называется электрической постоянной, а численный множитель 4π вводится для рационализации:

$$F=\frac{1}{4\pi\varepsilon_0}\frac{q_1q_2}{r^2}. \tag{11}$$

Из формулы (11) видно, что размерность ε_0 есть Кл²/(Н·м²), т.е. $M^{-1}L^{-3}T^4I^2$. Численное значение электрической постоянной ε_0 определяется на опыте. Его можно найти, измеряя, например, силу взаимодействия точечных зарядов известной величины, находящихся на известном расстоянии друг от друга. Измерения дают для ε_0 следующее значение:

$$\varepsilon_0 = 8{,}854 \cdot 10^{-12} \quad \text{Кл}^2/(\text{Н}\cdot\text{м}^2). \tag{12}$$

Установить соотношение между единицами заряда (или тока) в системах СГСЭ (или гауссовой) и СИ можно следующим образом. Пусть по параллельным проводам, находящимся на расстоянии 1 м друг от друга, текут токи по 1 А. Тогда действующая на 1 м длины провода сила равна $2\cdot10^{-7}$ Н, т. е. $2\cdot10^{-2}$ дин. Формула (7) для вычисления этой силы в гауссовой системе единиц имеет вид

$$F = \frac{1}{c^2}\,\frac{2I^2l}{r}.$$

Подставляя в левую часть этой формулы $F=2\cdot10^{-2}$ дин, а в правую $l=r=100$ см, $c\approx3\cdot10^{10}$ см/с, найдем, что такую силу магнитного взаимодействия проводников обеспечивает ток, численно равный $0{,}1c$ единиц тока в системе СГСЭ. Итак, один ампер равен $0{,}1c$ единиц тока СГСЭ:

$$\frac{1\ \text{А}}{1\ \text{ед. тока СГСЭ}} = 0{,}1c = 2{,}9979246\cdot10^{9} \approx 3\cdot10^{9}.$$

Соотношение между единицами тока в гауссовой системе и системе СИ выражается через определяемую на опыте постоянную c.

Таким же является и соотношение между единицами заряда в этих системах: 1 Кл$\approx3\cdot10^9$ ед. заряда СГСЭ. Зная это соотношение, можно вычислить величину электрической постоянной ε_0. Два точечных заряда по 1 кулону, находящиеся на расстоянии 1 м друг от друга, взаимодействуют с силой, равной $1/(4\pi\varepsilon_0)$ ньютонов. В гауссовой системе единиц эта же сила равна

$$F = \frac{q^2}{r^2} = \frac{(3\cdot10^9)^2}{100^2}\ \text{дин} = 9\cdot10^9\ \text{Н}.$$

Итак, численное значение $1/(4\pi\varepsilon_0)$ равно $9\cdot10^9$. Поэтому

$$\varepsilon_0 = \frac{1}{4\pi\cdot9\cdot10^9}\,\frac{\text{Кл}^2}{\text{Н}\cdot\text{м}^2} = 8{,}854\cdot10^{-12}\,\frac{\text{Кл}^2}{\text{Н}\cdot\text{м}^2}.$$

Нетрудно видеть, что фактически мы не теоретически «вычислили» ε_0, а всего лишь выразили ее через другую экспериментально определяемую постоянную c. Это означает, что в конечном счете электрическая

постоянная все-таки определяется экспериментально. Этим она принципиально отличается от магнитной постоянной μ_0, значение которой задается (а не измеряется!) в связи с определением ампера. Такое различие между ε_0 и μ_0 связано с тем, что в системе СИ введение единицы заряда основано на магнитном взаимодействии токов, а не на электростатическом взаимодействии зарядов.

Связь между экспериментально определяемыми постоянными ε_0 и c может быть представлена в несколько иной форме. Для этого сравним выражения для сил электростатического взаимодействия зарядов и магнитного взаимодействия токов, записанные в системах СИ и Гаусса:

$$F_э = \frac{1}{4\pi\varepsilon_0}\frac{q^2}{r^2}, \qquad F_э = \frac{q^2}{r^2}, \tag{13}$$

$$F_м = \frac{\mu_0}{4\pi}\frac{2I^2l}{r}, \qquad F_м = \frac{1}{c^2}\frac{2I^2l}{r}. \tag{14}$$

Безразмерное отношение $F_э/F_м$ должно быть одинаковым в обеих системах единиц. Составляя отношение правых частей выражений (13), (14) и приравнивая значения этого отношения в СИ и гауссовой системе, можно убедиться, что

$$\frac{1}{c^2} = \varepsilon_0\mu_0. \tag{15}$$

Как уже отмечалось, запись некоторых формул в гауссовой системе и в СИ отличается не только размерными коэффициентами, но и численными множителями, появление которых связано с рационализацией системы единиц СИ. Ниже приводится таблица, в которой сопоставляется запись наиболее важных формул электромагнетизма в этих системах единиц.

В заключение отметим, что наряду с единицами, входящими в системы, нередко используются так называемые внесистемные единицы. Ниже приводятся некоторые широко используемые внесистемные единиц.

1. Сила: 1 килограмм-сила (кгс)=9,81 Н.
2. Давление: 1 нормальная атмосфера (атм)=760 мм рт. ст.= $=1{,}013\cdot10^5$ Н/м².
3. Работа и энергия: 1 килограмм-сила-метр (кгс·м)=9,81 Дж;
   1 калория (кал)=4,185 Дж;
   1 электрон-вольт (эВ)=$1{,}6\cdot10^{-19}$ Дж.
4. Мощность: 1 лошадиная сила (л.с.)=736 Вт.

Таблица. Основные формулы электромагнетизма в системах СИ и гауссовой

Название	Гауссова система	СИ
1. Закон Кулона	$F=\frac{q_1q_2}{r^2}$	$F=\frac{1}{4\pi\varepsilon_0}\frac{q_1q_2}{r^2}$
2. Напряженность электрического поля (определение)	$E=\frac{F}{q}$	
3. Напряженность поля точечного заряда	$E=\frac{q}{r^2}$	$E=\frac{1}{4\pi\varepsilon_0}\frac{q}{r^2}$
4. Напряженность поля вблизи поверхности проводника	$E=4\pi\sigma$	$E=\frac{\sigma}{\varepsilon_0}$
5. Поток напряженности электрического поля через поверхность	$N=\sum E_n\Delta S$	
6. Теорема Гаусса	$N=4\pi q$	$N=\frac{q}{\varepsilon_0}$
7. Потенциал электрического поля (определение)	$\varphi=\frac{A}{q}$	
8. Потенциал поля точечного заряда	$\varphi=\frac{q}{r}$	$\varphi=\frac{1}{4\pi\varepsilon_0}\frac{q}{r}$
9. Связь между E и φ	$E_l=-\frac{\Delta\varphi}{\Delta l}$	
10. Емкость (определение)	$C=\frac{q}{U}$	
11. Емкость плоского конденсатора с диэлектриком	$C=\frac{\varepsilon S}{4\pi d}$	$C=\frac{\varepsilon_0\varepsilon S}{d}$
12. Энергия системы зарядов	$W=\frac{1}{2}\sum_i q_i\varphi_i$	
13. Энергия конденсатора	$W=\frac{1}{2}CU^2$	
14. Плотность энергии электрического поля	$w=\frac{E^2}{8\pi}$	$w=\frac{\varepsilon_0E^2}{2}$

Продолжение

Название	Гауссова система	СИ
15. Сила тока	$I=\frac{dq}{dt}$	
16. Закон Ома	$I=\frac{U}{R}$	
17. Закон Джоуля—Ленца	$Q=I^2Rt$	
18. Сила взаимодействия двух параллельных токов в вакууме	$F=\frac{1}{c^2}\frac{2I_1I_2l}{r}$	$F=\frac{\mu_0}{4\pi}\frac{2I_1I_2l}{r}$
19. Закон Био—Савара—Лапласа	$\Delta B=\frac{1}{c}\frac{I\,\Delta l\sin\alpha}{r^2}$	$\Delta B=\frac{\mu_0}{4\pi}\frac{I\,\Delta l\sin\alpha}{r^2}$
20. Индукция магнитного поля прямого тока	$B=\frac{1}{c}\frac{2I}{r}$	$B=\frac{\mu_0}{2\pi}\frac{I}{r}$
21. Индукция магнитного поля в центре кругового тока	$B=\frac{1}{c}\frac{2\pi I}{r}$	$B=\frac{\mu_0 I}{2r}$
22. Индукция поля в соленоиде	$B=\frac{4\pi}{c}nI$	$B=\mu_0 In$
23. Циркуляция вектора **B**	$\sum B_l\Delta l=\frac{4\pi}{c}I$	$\sum B_l\Delta l=\mu_0 I$
24. Закон Ампера	$F=\frac{1}{c}IBl\sin\alpha$	$F=IBl\sin\alpha$
25. Сила Лоренца	$F=\frac{1}{c}q\boldsymbol{v}\times\boldsymbol{B}$	$F=q\boldsymbol{v}\times\boldsymbol{B}$
26. Поток магнитной индукции (определение)	$\Phi=\sum B_n\Delta S$	
27. Закон электромагнитной индукции	$\mathscr{E}=-\frac{1}{c}\frac{d\Phi}{dt}$	$\mathscr{E}=-\frac{d\Phi}{dt}$
28. Индуктивность (определение)	$\Phi=\frac{1}{c}LI$	$\Phi=LI$
29. Индуктивность соленоида	$L=4\pi n^2V$	$L=\mu_0 n^2V$
30. Магнитная энергия тока (энергия магнитного поля)	$W=\frac{1}{c^2}\frac{LI^2}{2}$	$W=\frac{LI^2}{2}$
31. Плотность энергии магнитного поля	$w=\frac{B^2}{8\pi}$	$w=\frac{B^2}{2\mu_0}$

СОДЕРЖАНИЕ

3. ЭЛЕКТРИЧЕСТВО И МАГНЕТИЗМ

4. КОЛЕБАНИЯ И ВОЛНЫ

5. ОПТИКА. ТЕОРИЯ ОТНОСИТЕЛЬНОСТИ

6. ВВЕДЕНИЕ В КВАНТОВУЮ ФИЗИКУ

1р. 20к.